AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

A. B. COBLE
UNIVERSITY OF ILLINOIS

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

G. C. EVANS
RICE INSTITUTE

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY
E. T. BELL
W. A. MANNING
HARRY BATEMAN
J. R. KLINE
E. P. LANE
HARRY LEVY
MARSTON MORSE

510.5
.002

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

510.5

VOLUME LII

1930

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

P14765

PRINTED BY THE J. H. FURST COMPANY
BALTIMORE, MD.

INDEX.

PAGE

TABLE OF CONTENTS.

---

1. *Statement of the problem.* We are to consider the problem of minimizing an integral of the form

$$J=\int_{t_0}^{T} F(x_1,\cdot\cdot\cdot,x_n,x_1',\cdot\cdot\cdot,x_n')\,dt=\int_{t_0}^{T} F(x,x')\,dt$$

in the class of all admissible curves

$$x^\alpha=x^\alpha(t) \qquad (t_0\leqq t\leqq T;\,\alpha=1,\cdot\cdot\cdot,n)$$

which join two fixed points x_0^α and X^α in space of n dimensions $(n>1)$. We assume the usual continuity and homogeneity properties for the integrand function F, and "admissible curves" are ordinary curves in the sense used by Bolza.* We shall make use of the well-known summation convention of

* *Vorlesungen*, § 25, p. 189.

tensor analysis, except that it will not be necessary to distinguish between contravariant and covariant indices. We have, for example, to deal with the "norm" $x'^a x'^a$ ($= \sum_{a=1}^{n} x'^a x'^a$) of a vector x'^a. It is important to recall that if the x'^a are the derivatives of functions x^a defining an admissible curve, then $x'^a x'^a$ never vanishes along that curve. We shall use the indices α and β to run from 1 to n, and the indices i and j to run from 1 to $n-1$.

2. *The classical necessary conditions for a minimum.* An admissible curve

$$E_0 : x^a = x^a(t)$$

minimizing the integral J must satisfy the following well-known necessary conditions:

(I_1) the *Euler equations:*

$$\frac{d}{dt}\,\frac{\partial F}{\partial x'^a} = \frac{\partial F}{\partial x^a} \tag{2.1}$$

between corners on E_0;

(I_2) the *corner conditions:*

$$\frac{\partial F(x, x'^-)}{\partial x'^a} = \frac{\partial F(x, x'^+)}{\partial x'^a} \tag{2.2}$$

at each corner; here x'^- and x'^+ denote the directions on E_0 preceding and following the corner, respectively;

(II) the *Weierstrass condition:*

$$\mathcal{E}(x, x', \bar{x}') \equiv \left[\frac{\partial F(x, \bar{x}')}{\partial x'^a} - \frac{\partial F(x, x')}{\partial x'^a}\right]\bar{x}'^a \geqq 0$$

for every (x, x') on E_0 and every direction $\bar{x}'$ different from x';

(III) the *Legendre condition:*

$$Q(x, x'; \eta) \equiv \eta^a \frac{\partial^2 F(x, x')}{\partial x'^a \partial x'^\beta}\,\eta^\beta \geqq 0$$

for every (x, x') on E_0 and every η whose components are not proportional to those of x';

(III_0) the function F_1 defined by the equations *

---

* Cf. Bliss, *Transactions of the American Mathematical Society*, Vol. 15 (1914), pp. 376, 378.

$$F_1 x'^{\alpha} x'^{\beta} = H^{\alpha\beta}$$

is not negative along E_0; here the matrix $H^{\alpha\beta}$ is the adjoint of $\partial^2 F/\partial x'^{\alpha}\partial x'^{\beta}$;

(IV) the *Jacobi condition:* that there shall be no pair of conjugate points on any arc of E_0 which is of class C' and has $F_1 \neq 0$ along it.

Geometrically interpreted, the Jacobi condition is that no arc of E_0 of class C' may contain a point of contact with the envelope, if existing, of the $(n-1)$-parameter family of extremal arcs passing through a fixed point of the arc of E_0 in question. If the equations of this family of extremal arcs are

$$x^{\alpha} = \psi^{\alpha}(t, a)$$

then the analytic formulation of the condition is that the determinant

$$\left| \partial\psi^{\alpha}/\partial a^{i} \qquad \partial\psi^{\alpha}/\partial t \right|$$

shall not vanish along the arc of E_0 in question, except at the fixed point through which all the extremals of the family pass. A formulation of the Jacobi condition for minimizing arcs of class C', in terms of the second variation and solutions of the Jacobi equations, has been given by Bliss.* A modification of Bliss' method is used in §§ 8-10 of this paper.

3. *Construction of families of extremaloids.* The equations (2. 1) and (2. 2) characterize the admissible curves called *simple extremaloids* by Tonelli.† We shall drop the prefix "simple," inasmuch as no other types of extremaloids appear in this paper.

If E_0 is an extremal arc

$$x^{\alpha} = x^{\alpha}(t) \qquad (t_1 \leqq t \leqq t_2)$$

along which the function F_1 does not vanish, then for each point x_0 and direction x_0' sufficiently near those on E_0 there is a unique extremal passing through x_0 in the direction x_0'. The equations of these extremals can be given in the form

$$x^{\alpha} = \Xi^{\alpha}(t, x_0, x_0')$$

where the functions Ξ^{α} and their first and second derivatives with respect to t are of class C' for the specified range of x_0 and x_0' and for t on an interval $t_1 - \delta \leqq t \leqq t_2 + \delta$, and where

$$\Xi^{\alpha}(t, x_0, x_0') = x^{\alpha}(t) \qquad (t_1 \leqq t \leqq t_2)$$

* "Jacobi's Condition for Problems of the Calculus of Variations in Parametric Form," *Transactions of the American Mathematical Society*, Vol. 17 (1916), p. 195.

† *Fondamenti di Calcolo delle Variazioni*, Vol. II, p. 189.

whenever x_0 and x_0' represent a point and corresponding direction on the original extremal E_0. The totality of these extremals constitutes a $(2n-2)$-parameter family

$$x^a = \Psi^a(t, a, b)$$

where a and b each represent $n-1$ parameters. The totality of extremals through a fixed point constitutes an $(n-1)$-parameter family

$$x^a = \psi^a(t, a).$$

The functions Ψ^a and ψ^a have the same differentiability properties as Ξ^a. Moreover the determinant

$$(3.1) \qquad \begin{vmatrix} \dfrac{\partial \Psi^a}{\partial a^i} & \dfrac{\partial \Psi^a}{\partial b^i} & \dfrac{\partial \Psi^a}{\partial t} & 0 \\ \dfrac{\partial^2 \Psi^a}{\partial t \partial a^i} & \dfrac{\partial^2 \Psi^a}{\partial t \partial b^i} & \dfrac{\partial^2 \Psi^a}{\partial t^2} & \dfrac{\partial \Psi^a}{\partial t} \end{vmatrix}$$

never vanishes for $t_1 - \delta \leqq t \leqq t_2 + \delta$, and the determinant

$$(3.2) \qquad \begin{vmatrix} \dfrac{\partial \psi^a}{\partial a^i} & \dfrac{\partial \psi^a}{\partial t} \end{vmatrix}$$

is not identically zero on any subinterval of $t_1 t_2$.*

We shall obtain corresponding imbedding theorems for extremaloids, by starting from a family of extremals containing an extremal arc of the given extremaloid E_0, and showing how to get past a corner.† The extension of the properties of the determinants (3.1) and (3.2) will come naturally in the later sections on the second variation and the Jacobi equations.

Suppose then that E_0 is an extremaloid joining the points x_0 and X and having corners at $x_1, \cdots, x_m$, and suppose that the function F_1 is different from zero along E_0, including both sides of corners. For definiteness and simplicity of notation, start from a family $x^a = \psi^a(t, a)$ containing the first extremal arc E_{01} of E_0 for $a^i = a_0^i$, $t_0 \leqq t \leqq t_1$. Let p^{a-} and p^{a+} repre-

* These statements are justified on the basis of the imbedding theorems for differential equations. Cf. Bolza, *Vorlesungen über Variationsrechnung*, Ch. IV; Bliss and Mason, *Transactions of the American Mathematical Society*, Vol. 9 (1908), pp. 443, 444; J. H. Taylor, *Bulletin of the American Mathematical Society*, Vol. 31 (1925), p. 257.

† Cf. the references to Caratheodory and Bolza already cited; also Sakellariou, "Sur les solutions discontinues du problème du calcul des variations dans l'espace à n dimensions," *Comptes Rendus du Congrés Internationale des Mathématiciens*, Strasbourg (1924), p. 351.

sent the direction cosines of the tangents to E_0 at the corner x_1, and suppose for convenience that $\psi^{a\prime}(t_1, a_0) = p^{a-}$.* The corner equations

$$(3.3) \qquad \frac{\partial F[\psi(t,a), q]}{\partial x'^a} - \frac{\partial F[\psi(t,a), \psi'(t,a)]}{\partial x'^a} = 0, \qquad q^a q^a - 1 = 0$$

have the initial solution $a^i = a_0{}^i$, $t = t_1$, $q^a = p^{a+}$. At this solution the functional determinant with respect to the $n+1$ variables (t, q) is

$$(3.4) \qquad \begin{vmatrix} \dfrac{\partial^2 F^+}{\partial x'^a \partial x'^\beta} & \dfrac{\partial^2 F^+}{\partial x'^a \partial x^\beta} p^{\beta-} - \dfrac{\partial F^-}{\partial x^a} \\ 2p^{\beta+} & 0 \end{vmatrix},$$

since the arc E_{01} satisfies the Euler equations (2.1). If we expand this determinant by the last row and column and make use of the identities †

$$p^{a+} F_1{}^+ p^{\beta+} = \text{adjoint of } \frac{\partial^2 F^+}{\partial x'^a \partial x'^\beta},$$
$$p^{a+} \frac{\partial^2 F^+}{\partial x'^a \partial x^\beta} = \frac{\partial F^+}{\partial x^\beta},$$

we find that the determinant (3.4) reduces to $2F_1{}^+ \Omega_0$, where

$$\Omega_0(x, p^-, p^+) = (\partial F^+/\partial x^a) p^{a-} - (\partial F^-/\partial x^a) p^{a+}.$$

Consequently, *if we make the additional assumption*

$$\Omega_0 \neq 0,$$

the equations (3.3) have unique solutions $t = t(a)$, $q^a = q^a(a)$, having a, t, q in sufficiently restricted neighborhoods of a_0, t_1, p^+ respectively, and these solutions are of class C' at least. Application of the imbedding theorems to the external arc E_{02} adjacent to E_{01} shows that through each point $x^a = \psi^a(t(a), a)$ there passes a unique extremal arc in the direction $x'^a = q^a(a)$, provided a^i is sufficiently near $a_0{}^i$.

If the function Ω_0 is different from zero at each of the corners on E_0, we obtain by this method (with proper choice of the parameter t on each sub-arc) *a family $x^a = \phi^a(t, a)$ of extremaloids, defined for $t_0 - \delta \leqq t \leqq T + \delta$, containing the original extremaloid E_0 for $a = a_0$, $t_0 \leqq t \leqq T$. The functions ϕ^a are continuous in all their arguments, and ϕ^a, $\partial\phi^a/\partial t$, and $\partial^2\phi^a/\partial t^2$ are of class C' in all their arguments except at the corners.* There is no change in the reasoning if we start from a family of extremals depending on a different number of parameters. We shall use the notation

---

* Primes will always denote differentiation with respect to t.

† These identities are consequences of the homogeneity condition on the integrand function F.

$$x^a = \Phi^a(t, a, b)$$

to represent the general $(2n-2)$-parameter family of extremaloids, containing E_0 for $a^i = a_0{}^i$, $b^i = b_0{}^i$. That we obtain the same set of curves, no matter which extremal arc of E_0 we start from, is a consequence of the fact, to be proved in § 10 for the case of extremaloids, that if the determinant

$$\begin{vmatrix} \dfrac{\partial \Phi^a}{\partial a^i} & \dfrac{\partial \Phi^a}{\partial b^i} & \dfrac{\partial \Phi^a}{\partial t} & 0 \\ \dfrac{\partial^2 \Phi^a}{\partial t \partial a^i} & \dfrac{\partial^2 \Phi^a}{\partial t \partial b^i} & \dfrac{\partial^2 \Phi^a}{\partial t^2} & \dfrac{\partial \Phi^a}{\partial t} \end{vmatrix}$$

is different from zero at one point of E_0 then it never vanishes.

It is sometimes convenient to have the corners on all the extremaloids of a family occur for the same values of the parameter t. This can be brought about by a transformation of the form

$$t = f(\bar{t}, a, b),$$

where f is continuous in all its arguments, f and $f' = \partial f / \partial \bar{t}$ are of class C' between corners, and $f' > 0$. Such a transformation multiplies the determinants

$$\begin{vmatrix} \dfrac{\partial \phi^a}{\partial a^i} & \dfrac{\partial \phi^a}{\partial t} \end{vmatrix}, \quad \begin{vmatrix} \dfrac{\partial \Phi^a(t)}{\partial a^i} & \dfrac{\partial \Phi^a(t)}{\partial b^i} & \dfrac{\partial \Phi^a(t)}{\partial t} & 0 \\ \dfrac{\partial \Phi^a(t_0)}{\partial a^i} & \dfrac{\partial \Phi^a(t_0)}{\partial b^i} & 0 & \dfrac{\partial \Phi^a(t_0)}{\partial t} \end{vmatrix},$$

by $f'(\bar{t})$ and $f'(\bar{t})f'(\bar{t}_0)$ respectively, and hence cannot affect the sign of either of these determinants. When the corners on all the extremaloids of the family occur for the same values of the parameter t, the partial derivatives of the functions Φ^a (or ϕ^a) with respect to the parameters a^i and b^i are continuous functions of t.

4. *A relation between the functions $\mathcal{E}$ and Ω_0, and a further necessary condition.* The relation sought is due to Dresden.* We temporarily select the arc length s as parameter along the extremaloid E_0. Then we find preceding a corner

$$d/ds\, \mathcal{E}\,[x(s), x'(s), p^+] = \Omega_0,$$

and following a corner

$$d/ds\, \mathcal{E}\,[x(s), x'(s), p^-] = -\Omega_0,$$

* *Transactions of the American Mathematical Society,* Vol. 9 (1908), p. 485.

where $p^{\alpha-}$ are the direction cosines of the tangent to E_0 preceding the corner, and $p^{\alpha+}$ those of the tangent following the corner. From the corner conditions (2.2) we obtain the equation

$$\mathcal{E}\ (x, p^-, p^+) = 0 = \mathcal{E}\ (x, p^+, p^-).$$

From this and the Weierstrass condition we readily obtain the following theorems.

If E_0 is an extremaloid along which $\mathcal{E} \geqq 0$, then $\Omega_0 \leqq 0$ at the corners.

If E_0 is an admissible arc minimizing the integral J, then $\Omega_0 \leqq 0$ at the corners on E_0.

5. *Additional properties of extremals and extremaloids.* The Legendre quadratic form $Q(x, x'; \eta)$ is said to be *positive regular* in case it is never negative and vanishes only when the η^α are proportional to x'^α. The form Q can be transformed by a real orthogonal transformation into a form $\bar{Q} = \lambda_\alpha \zeta^\alpha \zeta^\alpha$ whose coefficients λ_α are the n roots of the characteristic equation of the matrix $\partial^2 F / \partial x'^\alpha \partial x'^\beta$.* The characteristic equation in this case has the form when expanded

$$(-1)^n \lambda^n + \cdot \cdot \cdot \cdot \cdot - F_1 x'^\alpha x'^\alpha \lambda = 0.$$

Consequently we have

LEMMA 5.1. *If the form Q is never negative and $F_1 > 0$, then Q is positive regular, and conversely, if Q is positive regular, then $F_1 > 0$.*

Since the roots λ_α of the characteristic equation are continuous functions of (x, x'), we have the further property:

LEMMA 5.2. *If the Legendre quadratic form Q is positive regular for (x, x') in a bounded closed set S, then Q is positive regular in a neighborhood of S.*

The Weierstrassian function $\mathcal{E}\ (x, x', \bar{x}')$ obviously vanishes whenever x' and $\bar{x}'$ represent the same direction, and we wish to secure a related function which shall vanish only when the function $\mathcal{E}$ vanishes for x' and $\bar{x}'$ in different directions. This has been done by Bliss for space problems as follows.† It is first shown that for every two directions p and q there is at least one direction p_1 orthogonal to p such that

$$q^\alpha = p^\alpha \cos \omega + p_1{}^\alpha \sin \omega \tag{5.1}$$

* Cf. Kowalewski, *Einführung in die Determinantentheorie* (1909), § 116; Dickson, *Modern Algebraic Theories*, p. 76, theorem 12.

† The "Weierstrass $\mathcal{E}$-Function for Problems of the Calculus of Variations in Space," *Transactions of the American Mathematical Society*, Vol. 15 (1914), p. 369.

where $0 \leqq \omega \leqq \pi$.* For every pair of orthogonal directions p and p_1 we define: a direction q depending also on ω by equation (5.1), and directions q_1 and q_2 depending also on τ by

$$q_1{}^a = p^a \cos\tau + p_1{}^a \sin\tau, \qquad q_2{}^a = -p^a \sin\tau + p_1{}^a \cos\tau, \qquad (0 \leqq \tau \leqq \omega).$$

Then we define the desired new function $\mathcal{E}_1$ by the equations

$$\mathcal{E}_1(x, p, p_1, \omega) = \mathcal{E}(x, p, q)/(1 - \cos\omega) \qquad (0 < \omega \leqq \pi),$$
$$\mathcal{E}_1(x, p, p_1, 0) = Q(x, p; p_1),$$

where Q is the Legendre quadratic form. It is easily shown that

$$\mathcal{E}_1(x, p, p_1, \omega) = Q(x, q_1{}^*; q_2{}^*) \tag{5.2}$$

where $q_1{}^*$ and $q_2{}^*$ correspond to a value $\tau = \tau^*$ between 0 and ω. Hence the function $\mathcal{E}_1$ is continuous in all its arguments when p_1 is orthogonal to p.†

We say that an extremal arc E is positively strong in case

$$\mathcal{E}_1(x, p, p_1, \omega) > 0$$

for every element (x, p) on E, p_1 orthogonal to p, and $0 \leqq \omega \leqq \pi$. Since such a set of points (x, p, p_1, ω) is closed, we have the property that:

LEMMA 5.3. *If an extremal arc E is positively strong, then there is a neighborhood of the elements (x, p) on E in which $\mathcal{E}_1(x, p, p_1, \omega) > 0$ for all directions p_1 orthogonal to p and $0 \leqq \omega \leqq \pi$.*

An additional proposition, obvious from the definition of the function $\mathcal{E}_1$, is the following:

LEMMA 5.4. *If $\mathcal{E}_1(x, p, p_1, \omega) > 0$ at an element (x, p) for all directions p_1 orthogonal to p and $0 \leqq \omega \leqq \pi$, then the quadratic form Q is positive regular at (x, p).*

An important theorem, due to Caratheodory for the plane case,‡ is as follows:

LEMMA 5.5. *If an extremal arc E ceases to be strong at a point x_1, but still has the quadratic form $Q(x_1, x_1'; \eta)$ regular, then there is an admissible*

---

* The components of p, q, etc., are supposed to represent direction cosines.

† The extra argument p_1 is essential because when $n > 2$ there is more than one direction orthogonal to p, and the limit of the ratio $\mathcal{E}(x, p, q)/(1 - \cos\omega)$ when q approaches p depends on the direction of approach.

‡ Cf. Bolza, *Variationsrechnung*, p. 387.

direction $\bar{p}$ *which with the direction* x_1' *of* E *at* x_1 *satisfies the corner conditions* (2.2).

To prove this, let $x^a = x^a(t)$ be equations of the arc E, and let $t = t_1$ correspond to the point x_1. Suppose for definiteness that E is positively strong for $t < t_1$. Let

$$\mathcal{E}(t, q) \equiv \mathcal{E}[x(t), x'(t), q], \; \mathcal{E}_1(t, p_1, \omega) \equiv \mathcal{E}_1[x(t), x'(t), p_1, \omega].$$

Then $\mathcal{E}_1(t, p_1, \omega) > 0$ for every $t < t_1, p_1$ orthogonal to $x'(t)$, and $0 \leqq \omega \leqq \pi$. Consequently $\mathcal{E}_1(t_1, p_1, \omega) \geqq 0$ for the same range of p_1 and ω. Since E ceases to be strong at x_1, there is a direction p_1 orthogonal to x_1', and ω such that $\mathcal{E}_1(t_1, p_1, \omega) = 0$. Moreover, $\omega \neq 0$, since Q is regular at (x_1, x_1'). Hence there is a direction $\bar{p}$ different from x_1', such that $\mathcal{E}(t_1, \bar{p}) = 0$. By Taylor's theorem we have

$$\begin{aligned} \mathcal{E}(t_1, q) &= \mathcal{E}(t_1, \bar{p}) + \frac{\partial \mathcal{E}(t_1, \bar{p})}{\partial q^a}(q^a - \bar{p}^a) + R \\ &= \left[\frac{\partial F(x_1, \bar{p})}{\partial x'} - \frac{\partial F(x_1, x_1')}{\partial x'^a}\right](q^a - \bar{p}^a) + R. \end{aligned}$$

Since $\mathcal{E}(t_1, q) \geqq 0$ for every q, normed or not, and since R is infinitesimal of the second order with respect to the differences $q^a - \bar{p}^a$, we must have

$$(\partial/\partial x'^a) F(x_1, \bar{p}) = (\partial/\partial x'^a) F(x_1, x_1').$$

We shall say that an *extremaloid* E_0 is *positively strong* in case each of its extremal arcs is positively strong between corners, and the function F_1 remains positive on both sides of the corners on E_0. This requirement implies that the Legendre quadratic form Q is positive regular along E_0 including both sides of corners.

Consider a point x_1 and a direction p such that there is a unique direction p^+ different from p, satisfying the corner equations

$$(\partial/\partial x'^a) F(x_1, p) = (\partial/\partial x'^a) F(x_1, p^+).$$

In this case we shall say that at x_1 *there is a unique continuation direction* p^+ *corresponding to* p. It is obvious that p is then the unique continuation direction corresponding to p^+.

Suppose that E_0 is a positively strong extremaloid, having $\Omega_0 \neq 0$ at the corners, and with a unique continuation direction corresponding to each corner direction. By § 4, $\Omega_0 < 0$ at each corner, and if we continue past a corner on any extremal composing E_0, the function $\mathcal{E}(x, x', q)$ becomes negative for some directions q. If we require that the "strong" property is to

be preserved wherever possible, then an extremaloid having these propertie may be said to be *uniquely determined by each of its elements* (x, x').

We can now state the following important theorem.

LEMMA 5.6. *Let E_0 be an extremaloid which is positively strong an which has $\Omega_0 \neq 0$ at each corner. Suppose also that the extremaloid E_0 i uniquely determined by each of its elements (x, x'). Let $x^a = \Phi^a(t, a, b)$ be equations of the $(2n-2)$ parameter family of extremaloids containin E_0 for $a^i = a_0{}^i$, $b^i = b_0{}^i$.*

Then there is a neighborhood of (a_0, b_0) in which every correspondin extremaloid of the family is positively strong.

The proof is rather complicated, and begins with the following pre liminary proposition.

Suppose that E is an extremal arc with the properties:

1) *E has direction p at a point x_1, and at x_1 there is a unique continua tion direction p^+ corresponding to p;*
2) *the quadratic form Q is positive regular at (x_1, p);*
3) $F_1(x_1, p^+) \neq 0$;
4) $\Omega_0(x_1, p, p^+) \neq 0$;
5) *the family $x^a = \Psi^a(t, a, b)$ of extremals contains E for $a^i = a_0{}^i$ $b^i = b_0{}^i$;*
6) *the parameter value t_1 corresponds to the point x_1 on E.*

Then there are neighborhoods $N(a_0, b_0)$ and $N(t_1)$ such that for eac (a, b) in this $N(a_0, b_0)$ there is one and only one set of functions $[t(a, b)$ $q^a(a, b)]$ having $t(a, b)$ in the $N(t_1)$ and the direction $q^a(a, b)$ differen from $\Psi'^a[t(a, b), a, b]$, and satisfying the corner equations

$$(5.3)\qquad \frac{\partial F[\Psi(t, a, b), q]}{\partial x'^a} - \frac{\partial F[\Psi(t, a, b), \Psi'(t, a, b)]}{\partial x'^a} = 0,$$

$$q^a q^a - 1 = 0.$$

For, suppose the last conclusion untrue. In constructing a family of extrema loids in § 3 we showed that there are always neighborhoods $N(a_0, b_0)$ an $N(t_1, p^+)$ in which there are unique functions $t(a, b)$, $q^a(a, b)$, satisfyin the corner conditions. Then for every $N(a_0, b_0, t_1)$ there will be a poin (a, b, t, q) such that:

1) (a, b, t) is in this $N(a_0, b_0, t_1)$;
2) (a, b, t, q) satisfies the corner equations (5.3);

3) the direction q is different from the direction $\Psi'(t, a, b)$;

4) (t, q) is not in the neighborhood $N(t_1, p^+)$.

The set composed of such points (a, b, t, q) has at least one limit point $(a_0, b_0, t_1, \bar{q})$. Since $\bar{q} \neq p^+$, we must have $\bar{q} = p$. Now at these points (a, b, t, q) we have

$$0 = \frac{\partial F(\Psi, q)}{\partial x'^{\alpha}} - \frac{\partial F(\Psi, \Psi')}{\partial x'^{\alpha}}$$

$$= \int_0^1 \frac{\partial^2 F[\Psi, \Psi' + u(q - \Psi')]}{\partial x'^{\alpha} \partial x'^{\beta}} (q^{\beta} - \Psi'^{\beta}) du,$$

and therefore

$$0 = \int_0^1 Q[\Psi, \Psi' + u(q - \Psi'); q - \Psi'] du.$$

Since the quadratic form Q is positive regular at (x_1, p) it remains positive regular in a neighborhood of (x_1, p), and since the directions q and Ψ' are different we have

$$\int_0^1 Q[\Psi, \Psi' + u(q - \Psi'); q - \Psi'] du > 0$$

when (a, b, t, q) is sufficiently near (a_0, b_0, t_1, p). Thus the desired contradiction is secured.

To complete the proof of lemma 5.6, suppose that the corners on E_0 correspond to parameter values $t_1, \cdots, t_m$. Then from the preliminary proposition we know that there exist neighborhoods

$$N(a_0, b_0), N(t_1), \cdots, N(t_m),$$

such that, for each (a, b) in $N(a_0, b_0)$ and each integer k there is only one parameter value $t_k(a, b)$ in $N(t_k)$ at which a corner is possible. Consider one of the extremal arcs E_{0k} of E_0. It has equations

$$x^{\alpha} = \Phi^{\alpha}(t, a_0, b_0) \qquad (t_k \leqq t \leqq t_{k+1}).$$

Let $\bar{t}_k > t_k$ be in $N(t_k)$, $\bar{t}_{k+1} < t_{k+1}$ be in $N(t_{k+1})$. Then by lemma 5.3, there is a neighborhood $N_k(a_0, b_0)$ contained in $N(a_0, b_0)$ such that for (a, b) in $N_k(a_0, b_0)$ the arc

$$E_k: x^{\alpha} = \Phi^{\alpha}(t, a, b) \qquad (\bar{t}_k \leqq t \leqq \bar{t}_{k+1})$$

is positively strong. If there is a parameter value t in $N(t_k)$ at which the extremal containing the arc E_k ceases to be strong, there is a corner possible, by lemma 5.5. Hence this $t = t_k(a, b)$. Thus we find that for each (a, b) in $N_k(a_0, b_0)$ the extremal arc

$$x^a = \Phi^a(t, a, b) \qquad [t_k(a, b) \leqq t \leqq t_{k+1}(a, b)]$$

is positively strong except at the ends. If the neighborhood $N_k(a_0, b_0)$ is taken sufficiently small, the function F_1 will remain positive at the ends of these extremal arcs. Since each extremaloid is composed of only a finite number of extremal arcs, we have proved the lemma.

GEOMETRIC TREATMENT OF THE JACOBI-CARATHEODORY CONDITION.

6. *The extension of the Jacobi condition. Let E_0 be an extremaloid joining the points x_0 and X and minimizing the integral J, and let*

$$x^a = \phi^a(t, a) \qquad (t_0 \leqq t \leqq T)$$

be equations of the $n-1$ parameter family of extremaloids passing through the point x_0 and containing E_0 for $a^i = a_0{}^i$. Then the determinant

$$D(t, a) = \left| \frac{\partial \phi^a}{\partial a^i} \quad \frac{\partial \phi^a}{\partial t} \right|$$

*does not vanish for $a^i = a_0{}^i$, $t_0 < t \leqq T$, with the possible exception of the parameter values $t_k + 0$ corresponding to the sides of corners toward the second end-point X.**

The proof is made by means of the envelope theorem in the usual way, since (with the exception noted in the statement of the condition) there is a neighborhood of the point of contact with the envelope in which corners need not appear.† The differentiation of the integrals is slightly simplified if the corners occur for fixed values of t on all the extremaloids of the family.‡

The proof requires the following assumptions:

1) $F_1 \neq 0$ *along* E_0;
2) $\Omega_0 \neq 0$ *at the corners on* E_0;
3) *at the first zero of* $D(t, a_0)$ following t_0 *the partial derivative* $\partial D/\partial t$ *does not vanish;*
4) *the enveloping curve of the one parameter family of extremaloids containing E_0 and determined in the usual way, has a regressive branch at its point of contact with E_0.*

For the discussion of the existence and properties of the envelope, the

* This exception is removed in the proof by means of the second variation in §§ 8-10.

† For the proof cf. Bolza, *Variationsrechnung*, pp. 336, 378, 610; Bliss and Mason, *Transactions of the American Mathematical Society*, Vol. 9 (1908), pp. 449-451.

‡ See the closing paragraphs of § 3.

function F is supposed to be of class C^{iv} at least, so that all the operations of differentiation required in the theory may surely be carried out.

7. *A new form of the Caratheodory condition. Let E_0 be an extremaloid joining the points x_0 and X and minimizing the integral J, and let $\phi^a(t,a)$, $D(t,a)$ have the same meaning as in* § 6. *Then $D(t,a_0)$ does not change sign at the corners on E_0.*

The proof to be given in this section depends on the following assumptions:

1) *the extremaloid E_0 is positively strong;*
2) *$\Omega_0 \neq 0$ at the corners on E_0;*
3) *E_0 is uniquely determined by each of its elements (x, x').*

As indicated in the closing paragraphs of § 3, we may suppose in making the proof that the corners on all extremaloids of the family occur for fixed values of t. Then the partial derivatives $\partial\phi^a/\partial a^i$ are all continuous. Now suppose that the determinant D changes sign at a corner x_k where $t = t_k$. Let

$$x^a = \phi^{a-}(t,a) \qquad (t_0 \leqq t \leqq t_k + \delta)$$

denote the equations of the parts of the extremaloids preceding x_k, but with their last extremal arcs extended slightly beyond the corner manifold. Let

$$x^a = x^{a+}(t) \qquad (t_k - \delta \leqq t \leqq T)$$

denote the equations of the part of E_0 following x_k, but with its first extremal arc extended backward past the corner x_k. The equations

$$\phi^{a-}(t,a) = x^{a+}(u) \tag{7.1}$$

have the initial solution $t = t_k$, $a = a_0$, $u = t_k$, at which their functional determinant with respect to (a, t) is $D(t_k^-, a_0) \neq 0$. Hence these equations have a unique solution $a^i = \bar{a}^i(u)$, $t = \bar{t}(u)$ near this initial solution, and the functions $\bar{a}^i(u)$, $\bar{t}(u)$, are of class C'. Their derivatives $\bar{a}'^i$, $\bar{t}'$, satisfy the relations

$$(\partial\phi^{a-}/\partial t)\bar{t}' + (\partial\phi^{a-}/\partial a^i)\bar{a}'^i = x'^{a+}. \tag{7.2}$$

The one parameter family of curves E_u made up of the two parts

$$\begin{aligned} x^a &= \phi^{a-}[t, \bar{a}(u)] \qquad && [t_0 \leqq t \leqq \bar{t}(u)], \\ x^a &= x^{a+}(t) && (u \leqq t \leqq T), \end{aligned}$$

are all continuous admissible curves joining the points x_0 and X. The value of the integral J taken along E_u is a function $J(u)$ which is differentiable. Its derivative turns out to be

$$J'(u) = \bar{a}'^i \int_{t_0}^{\bar{t}(u)} \left(\frac{\partial F}{\partial x^a} \frac{\partial \phi^{a-}}{\partial a^i} + \frac{\partial F}{\partial x'^a} \frac{\partial^2 \phi^{a-}}{\partial a^i \partial t} \right) dt + F^- \bar{t}' - F^+,$$

where the arguments of F^- are $\phi^{a-}[\bar{t}(u), \bar{a}(u)]$, $\phi'^{a-}[\bar{t}(u), \bar{a}(u)]$, and those of F^+ are $x^{a+}(u)$, $x'^{a+}(u)$. By means of integration by parts, the Euler equations (2.1) with the corner conditions (2.2), the homogeneity condition, and the relations (7.1) and (7.2), this expression reduces to

$$\begin{aligned} J'(u) &= \partial F^-/\partial x'^a [(\partial \phi^{a-}/\partial a^i) \bar{a}'^i + (\partial \phi^{a-}/\partial t) \bar{t}'] - (\partial F^+/\partial x'^a) x'^{a+} \\ &= - x'^{a+} (\partial F^+/\partial x'^a - \partial F^-/\partial x'^a) \\ &= - \mathcal{E}(\phi^{a-}, \phi'^{a-}, x'^{a+}). \end{aligned}$$

If we solve equations (7.2) for the derivative $\bar{t}'$ at the value $u = t_k$, we find

$$\bar{t}'(t_k) = D(t_k + 0, a_0)/D(t_k - 0, a_0) < 0.$$

Consequently, for $u > t_k$ but near t_k we have $\bar{t}(u) < t_k$, and the point $\phi^{a-}[\bar{t}(u), \bar{a}(u)]$ precedes the corner manifold on each extremaloid. Then by lemma 5.6 $\mathcal{E} > 0$ and hence $J'(u) < 0$ for u sufficiently near t_k, and hence E_0 could not minimize the integral J.

THE SECOND VARIATION.*

8. *The extension of the Jacobi equations. Properties of their solutions.* If the extremaloid E_0 minimizes the integral J, then the second variation

$$J_2 \equiv \int_{t_0}^{T} 2\omega(t, \xi, \xi') \, dt$$

is greater than or equal to zero for all admissible variations $\xi^a(t)$ such that $\xi^a(t_0) = \xi^a(T) = 0$. The quadratic form ω is defined by the equation

$$2\omega(t, \xi, \xi') = \xi^a \frac{\partial^2 F}{\partial x^a \partial x^\beta} \xi^\beta + 2\xi^a \frac{\partial^2 F}{\partial x^a \partial x'^\beta} \xi'^\beta + \xi'^a \frac{\partial^2 F}{\partial x'^a \partial x'^\beta} \xi'^\beta.$$

Here and in the remainder of the discussion of the second variation, the arguments of the derivatives of F are always the functions $x^a(t)$, $x'^a(t)$ defining the extremaloid E_0. Admissible variations $\xi^a(t)$ are continuous on $(t_0 T)$ and have continuous first derivatives except at a finite number of points where one or more of the derivatives may have a finite jump. It will be con-

* The ensuing treatment of the second variation employs the elegant methods originated by Bliss. However, it is necessary to depart in a minor way from the methods for the parametric problem expounded in Bliss' paper, "Jacobi's Condition for Problems of the Calculus of Variations in Parametric Form," *Transactions of the American Mathematical Society*, Vol. 17 (1916), p. 195.

venient to say that such functions are of class D'. *It is assumed throughout that $F_1 \neq 0$ along the extremaloid E_0, and that $\Omega_0 \neq 0$ at the corners.*

If ξ is a variation giving to the second variation J_2 its minimum value zero, the functions ξ^a must satisfy the equations

$$\frac{\partial\omega}{\partial\xi'^a} = \int_{t_0}^{t} \frac{\partial\omega}{\partial\xi^a}\, dt + c_a \qquad (t_0 \leqq t \leqq T), \tag{8.1}$$

where the c_a are properly chosen constants. These equations are equivalent to the differential equations

$$\frac{d}{dt}\frac{\partial\omega}{\partial\xi'^a} = \frac{\partial\omega}{\partial\xi^a} \tag{8.2}$$

holding between corners of E_0 and ξ, and the corner conditions

$$\frac{\partial\omega(t^-, \xi, \xi'^-)}{\partial\xi'^a} = \frac{\partial\omega(t^+, \xi, \xi'^+)}{\partial\xi'^a}. \tag{8.3}$$

These equations may be obtained by applying to the integral J_2 the Euler equations (2.1) and the corner conditions (2.2), which are valid even if the integrand function 2ω is discontinuous in t.

The equations (8.2) are the Jacobi differential equations. They are not independent, but in fact are satisfied between corners on E_0 by every set of functions of the form $\rho x'^a$ where ρ is a differentiable function of t. The equations (8.1) are to be looked upon as the extension of the Jacobi equations for the case of discontinuous solutions. For these equations we have the following property.

LEMMA 8.1. *Every set of functions of the form $\xi^a = \rho x'^a$ is a solution of equations (8.1), provided ρ is a function of t of class D' on (t_0T), and $\rho(t_1) = \cdots = \rho(t_m) = 0$.*

The parameter values $t_1, \cdots, t_m$, as before, correspond to the corners on E_0. To prove the lemma, we note first that

$$\begin{aligned}\frac{\partial\omega(t, \rho x', \rho x'' + \rho' x')}{\partial\xi^{\beta\prime}} &= \rho x'^a \frac{\partial^2 F}{\partial x^a \partial x'^\beta} + \rho x''^a \frac{\partial^2 F}{\partial x'^a \partial x'^\beta} + \rho' x'^a \frac{\partial^2 F}{\partial x'^a \partial x'^\beta} \\ &= \rho \frac{d}{dt}\frac{\partial F}{\partial x'^\beta} = \rho \frac{\partial F}{\partial x^\beta}\end{aligned} \tag{8.4}$$

on account of the homogeneity property of F and the Euler equations (2.1). Similarly

$$\begin{aligned}\frac{\partial\omega(t, \rho x', \rho x'' + \rho' x')}{\partial\xi^\beta} &= \rho x'^a \frac{\partial^2 F}{\partial x^a \partial x^\beta} + \rho x''^a \frac{\partial^2 F}{\partial x'^a \partial x^\beta} + \rho' x'^a \frac{\partial^2 F}{\partial x'^a \partial x^\beta} \\ &= \rho \frac{d}{dt}\frac{\partial F}{\partial x^\beta} + \rho' \frac{\partial F}{\partial x^\beta} = \frac{d}{dt}\left(\rho \frac{\partial F}{\partial x^\beta}\right).\end{aligned}$$

Moreover, the corner conditions (8. 3) are satisfied, since ρ is zero at $t_1, \cdots, t_m$, and ρ' disappears from the final form of $\partial\omega/\partial\xi'^\beta$ in (8. 4).

We shall say that a solution ξ^a of equations (8. 2) on an interval (t_{k-1}, t_k) is *normal* on that interval in case the functions ξ^a are of class C' and satisfy the equation $x'^a\xi'^a = 0$.* Using the assumption already made that $F_1 \neq 0$ along E_0, it is readily proved that for every normal solution the functions ξ^a are of *class* C'' at least. Solutions ξ^a normal on an interval (t_{k-1}, t_k) have the property that $\xi^a(\tau) = \xi'^a(\tau) = 0$ *at a point* τ *only if the functions* ξ^a *are identically zero on the interval.*† Moreover, if we start with an arbitrary solution ξ^a of equations (8. 2) which is of class C' on (t_{k-1}, t_k), we can secure *an associated normal solution* $\xi_N{}^a$ of the form $\xi_N{}^a = \xi^a - \rho x'^a$, where ρ is of class C'. Such a normal solution is *uniquely determined* by prescribing the value of ρ at a single point. This follows from the fact that the condition for $\xi^a - \rho x'^a$ to be normal yields a first order differential equation for ρ. With these facts in mind we can prove the important

LEMMA 8. 2. *If* ξ^a *is a solution of equations* (8. 1) *having*

$$\xi^a(\tau) = \gamma x'^a(\tau),\ \xi'^a(\tau) = \gamma x''^a(\tau) + \delta x'^a(\tau),$$

at a single point τ, *then* $\xi^a = \rho x'^a$ *on the whole interval* (t_0T), *where the function* ρ *is of class* D' *and vanishes at the corner parameter values* $t_1, \cdots, t_m$.

Consider an interval containing the point τ, on which the given solution ξ^a is of class C', and consider the associated normal solution $\xi_N{}^a = \xi^a - \rho x'^a$ having $\rho(\tau) = \gamma$ and hence $\xi_N{}^a(\tau) = 0$. The condition $x'^a\xi_N{}'^a = 0$, applied at the point τ, gives $\rho'(\tau) = \delta$, and hence $\xi_N{}'^a(\tau) = 0$. Consequently the normal solution $\xi_N{}^a$ is identically zero and $\xi^a = \rho x'^a$ on the interval in question. Consider next a point τ of discontinuity of ξ'^a, different from $t_1, \cdots, t_m$. At such a point τ the corner conditions (8. 3) reduce to

$$(\partial^2 F/\partial x'^a \partial x'^\beta)(\xi'^{\beta -} - \xi'^{\beta +}) = 0.$$

Since $F_1 \neq 0$, the matrix of coefficients has rank $n - 1$, and hence $\xi'^{\beta +}$ differs from $\xi'^{\beta -}$ by a multiple of x'^β. Then by the part of the lemma already proved, ξ^a has the form $\rho x'^a$ on the adjacent interval on which ξ^a is of class C'. Finally, suppose that the hypothesis of the lemma holds at $\tau = t_k^-$, corresponding to one side of one of the corners on E_0. By means of the Euler

* Alternative equations equally applicable to our needs would be: $x'^a\xi'^a =$ constant, or, $x'^a\xi^a =$ constant, or, $x'^a\xi^a =$ linear function of t.

† The proof of this is similar to the proof of Lemma 1, page 199, in the paper by Bliss already cited.

2

equations (2.1) and the homogeneity condition on the function F, the corner conditions (8.3) in this case are easily reduced to

$$(8.5) \qquad \gamma(\partial F^-/\partial x^\alpha) - (\partial^2 F^+/\partial x'^\alpha \partial x'^\beta)\xi'^{\beta+} - \gamma(\partial^2 F^+/\partial x'^\alpha \partial x^\beta)x'^{\beta-} = 0.$$

If we multiply these equations by $x'^{\alpha+}$ and sum, we find that the middle term drops out, and

$$\gamma[x'^{\alpha+}(\partial F^-/\partial x^\alpha) - (\partial F^+/\partial x^\beta)x'^{\beta-}] = -\gamma\Omega_0 = 0.$$

Since we have assumed $\Omega_0 \neq 0$ at the corners on E_0, we must have $\gamma = 0$. Substituting this in equations (8.5), we find that $\xi'^{\beta+}$ must be a multiple of $x'^{\beta+}$, so that if the hypothesis of the lemma holds on one side of a corner, it must hold on the other side also.

9. *A new definition of conjugate point, and the corresponding form of the Jacobi-Caratheodory condition.* A point x is defined to be *conjugate* to a point x_0 on an extremaloid E_0 in case it corresponds to a parameter value $\tau \neq t_0$ such that there exist constants γ, δ, ϵ, and an admissible solution ξ^α of the Jacobi equations (8.1) with the properties:

1) ξ^α is not of the form $\rho x'^\alpha$;
2) $\epsilon\xi^\alpha(t_0) = 0$, and $\epsilon\xi^\alpha(\tau) = \gamma x'^\alpha(\tau^-) + \delta x'^\alpha(\tau^+)$;
3) γ, δ, ϵ are not all three zero, and $\gamma\delta \geqq 0$.

Here $x'^\alpha(\tau^-)$ and $x'^\alpha(\tau^+)$ refer respectively to the left hand and right hand derivatives of x^α at τ. The special case where γ and δ may both be taken as zero is the only one of interest for a minimizing arc without corners. In case $\epsilon = 0$, it is obvious, since $x'^\alpha x'^\alpha \neq 0$, that $\gamma\delta > 0$ and x is at a corner which is a cusp on E_0.

THE JACOBI-CARATHEODORY CONDITION. *If E_0 is an admissible arc joining the points x_0 and X and minimizing the integral J, and if $F_1 \neq 0$ on E_0 and $\Omega_0 \neq 0$ at the corners on E_0, then there can be no point x conjugate to x_0 on E_0 between x_0 and X.*

Corollary. A minimizing arc can contain no cusp.

The hypotheses imply that E_0 is an extremaloid whose extremal arcs are each of class C''. Moreover, $\Omega_0 < 0$ at the corners, by § 4. Suppose x is conjugate to x_0 and lies between x_0 and X, and let ξ^α be a solution of the Jacobi equations defining the conjugate point x. Let

$$\begin{aligned} \eta^\alpha &= \epsilon\xi^\alpha - \rho x'^\alpha \qquad && (t_0 \leqq t \leqq \tau), \\ &= \rho x'^\alpha && (\tau \leqq t \leqq T), \end{aligned}$$

where the function ρ vanishes at $t_0, t_1, \cdots, t_m, T$, except that at $t = \tau$, $\rho(\tau^-) = \gamma$, $\rho(\tau^+) = \delta$. We suppose that ρ is of class D' on $(t_0\,\tau)$ and on $(\tau\,T)$. Then η is an admissible variation, vanishing at t_0 and T, and satisfying the Jacobi equations (8.1) on the intervals $(t_0\,\tau)$ and $(\tau\,T)$. However, η cannot satisfy those equations on the whole interval $(t_0\,T)$, by lemma 8.2, since if $\epsilon \neq 0$, η^a has the form $\rho x'^a$ on only part of the interval, and if $\epsilon = 0$, the parameter value τ corresponds to a corner at which ρ does not vanish. Consequently $J_2(\eta)$ cannot be a minimum. We shall show that $J_2(\eta) \leqq 0$, and hence that E_0 cannot minimize the integral J if E_0 contains a conjugate point x as assumed.

On an interval (t_{k-1}, t_k) not containing τ, the usual manipulations give [a]

$$\int_{t_{k-1}}^{t_k} 2\omega(t, \eta, \eta')\,dt = \eta^a \left.\frac{\partial\omega(t, \eta, \eta')}{\partial\xi'^a}\right|_{t_{k-1}}^{t_k}.$$

Since η^a vanishes at t_0 and T, and $\partial\omega/\partial\xi'^a$ is continuous except at $t = \tau$, we find

$$\begin{aligned} J_2(\eta) = \int_{t_0}^{T} 2\omega(t, \eta, \eta')\,dt &= \eta^a \left.\frac{\partial\omega(t, \eta, \eta')}{\partial\xi'^a}\right|_{\tau+}^{\tau-} \\ &= \eta^a\Big[\epsilon\,\frac{\partial\omega(\tau, \xi, \xi')^-}{\partial\xi'^a} - \frac{\partial\omega(\tau, \rho x', \rho x'' + \rho' x')^-}{\partial\xi'^a} \\ &\qquad - \frac{\partial\omega(\tau, \rho x', \rho x'' + \rho' x')^+}{\partial\xi'^a}\Big]. \end{aligned}$$

Now from equation (8.4) we have

$$(\partial/\partial\xi'^a)\,\omega(\tau, \rho x', \rho x'' + \rho' x') = \rho(\partial F/\partial x^a).$$

Also $\partial\omega/\partial\xi'^a$ is continuous on the solution ξ. Hence

$$\begin{aligned} J_2(\eta) &= \delta x'^{a+}[\epsilon(\partial/\partial\xi'^a)\omega(\tau, \xi, \xi')^+ - \gamma(\partial F^-/\partial x^a)] \\ &\quad - (\epsilon\xi^a - \gamma x'^{a-})\,\delta(\partial F^+/\partial x^a) \\ &= \delta\epsilon(\partial F^+/\partial x^\beta)\,\xi^\beta - \gamma\delta[x'^{a+}(\partial F^-/\partial x^a) - x'^{a-}(\partial F^+/\partial x^a)] \\ &\quad - \delta\epsilon\xi^a(\partial F^+/\partial x^a) \\ &= \begin{cases} 0 \text{ if } x \text{ is not at a corner,} \\ \gamma\delta\Omega_0 \leqq 0 \text{ if } x \text{ is at a corner.} \end{cases} \end{aligned}$$

10. *The relation between the two forms of the Jacobi-Caratheodory condition.* Associated with a set of $2n-2$ solutions

$$\xi_i{}^a, \eta_j{}^a \qquad (i, j = 1, \cdots, n-1),$$

[a] Cf. Bliss, *loc. cit.*, p. 200.

of the Jacobi equations (8.1) are two important determinants

$$\mathfrak{W}(t) = \begin{vmatrix} \xi_i{}^a & \eta_j{}^a & x'^a & 0 \\ \xi'_i{}^\beta & \eta'_j{}^\beta & x''^\beta & x'^\beta \end{vmatrix},$$

$$\Theta(t, t_0) = \begin{vmatrix} \xi_i{}^a(t) & \eta_j{}^a(t) & x'^a(t) & 0 \\ \xi_i{}^\beta(t_0) & \eta_j{}^\beta(t_0) & 0 & x'^\beta(t_0) \end{vmatrix},$$

and associated with a set of $n-1$ solutions is the determinant

$$D(t) = | \xi_i{}^a \quad x'^a |.$$

We shall proceed to develop the relations between these determinants, the points conjugate to x_0, and the families of extremaloids discussed in § 3.

LEMMA 10.1. *The determination $\mathfrak{W}(t)$ either vanishes identically or is never zero on the interval (t_0, T).*

To prove this, suppose $\mathfrak{W}(\tau) = 0$. Then there exist $2n$ constants a^i, b^i, γ, δ, not all zero, satisfying the equations

$$\begin{aligned} a^i\xi_i{}^a + b^i\eta_i{}^a + \gamma x'^a \qquad &= 0 \\ a^i\xi_i'^a + b^i\eta_i'^a + \gamma x''^a + \delta x'^a &= 0 \end{aligned}$$

for $t = \tau$. Moreover, the constants a^i, b^i, are not all zero, since if they were, we would have $\gamma = \delta = 0$ also. The solution $\xi^a = a^i\xi_i{}^a + b^i\eta_i{}^a$ of the Jacobi equations has $\xi^a(\tau) = -\gamma x'^a(\tau)$, $\xi'^a(\tau) = -\gamma x''^a(\tau) - \delta x'^a(\tau)$. Hence by lemma 8.2, ξ^a has the form $\rho x'^a$ on (t_0, T) and therefore $\mathfrak{W}(t)$ vanishes identically.

LEMMA 10.2. *If the determinant $\mathfrak{W}$ does not vanish, then every solution of the Jacobi equations* (8.1) *is expressible uniquely in the form*

$$\xi^a = a^i\xi_i{}^a + b^i\eta_i{}^a + \rho x'^a,$$

where the a^i, b^i, are constants, and ρ is a function of t of class D' vanishing at $t_1, \cdots, t_m$.

For at an arbitrary point τ of (t_0, T) the equations

$$\begin{aligned} \xi^a &= a^i\xi_i{}^a + b^i\eta_i{}^a + \gamma x'^a \\ \xi'^a &= a^i\xi_i'^a + b^i\eta_i'^a + \gamma x''^a + \delta x'^a \end{aligned}$$

have unique solutions for the $2n$ unknowns a^i, b^i, γ, δ. Then

$$\eta^a = \xi^a - a^i\xi_i{}^a - b^i\eta_i{}^a$$

is a solution of the Jacobi equations having $\eta^a(\tau) = \gamma x'^a(\tau)$, $\eta'^a(\tau) =$

$\gamma x''^a(\tau) + \delta x'^a(\tau)$, and consequently η^a has the form $\rho x'^a$ on the whole interval (t_0, T), by lemma 8.2.

LEMMA 10.3. *The function $\Theta(t, t_0)$ can be expressed in the form*

$$\Theta(t, t_0) = (t - t_0)^{n-1}\lambda(t, t_0),$$

where λ is continuous at $t = t_0$, and $\lambda(t_0, t_0) = \pm \mathfrak{W}(t_0)$.

Here it is understood that the parameter value t_0 does not correspond to a corner on E_0. Consequently the proof given in Bliss's paper * for his theorem 7 is essentially applicable to this lemma.

LEMMA 10.4. *The function $\Theta(t, t_0)$ vanishes identically if and only if $\mathfrak{W}(t)$ vanishes.*

That the condition $\mathfrak{W}(t) = 0$ is necessary follows at once from the lemmas 10.3 and 10.1. To show $\mathfrak{W}(t) = 0$ is sufficient, we have from the proof of lemma 10.1 that there exist constants a^i, b^i not all zero, and a function $\rho(t)$, such that $a^i\xi_i{}^a + b^i\eta_i{}^a + \rho x'^a \equiv 0$. It is an obvious consequence of this that the determinant $\Theta(t, t_0)$ vanishes identically in t and t_0.

LEMMA 10.5. *If the determinant $\Theta(t, t_0)$ does not vanish identically, then the points x conjugate to x_0 on the extremaloid E_0 correspond to the parameter values $t = \tau \neq t_0$ at which Θ vanishes or changes sign.*

By lemmas 10.2 and 10.4, a solution ξ^a of the Jacobi equations determining the conjugate point x is expressible in the form $\xi^a = a^i\xi_i{}^a + b^i\eta_i{}^a + \rho x'^a$, where the constants a^i, b^i, are not all zero, and the function ρ vanishes at the parameter values $t_1, \cdots, t_m$, corresponding to the corners. From the conditions $\epsilon\xi^a(\tau) = \gamma x'^a(\tau^-) + \delta x'^a(\tau^+)$, $\epsilon\xi^a(t_0) = 0$, we have the system of linear equations

$$\begin{gathered}\epsilon a^i\xi_i{}^a(\tau) + \epsilon b^i\eta_i{}^a(\tau) + \epsilon\rho(\tau)x'^a(\tau) - \gamma x'^a(\tau^-) - \delta x'^a(\tau^+) = 0\\ \epsilon a^i\xi_i{}^a(t_0) + \epsilon b^i\eta_i{}^a(t_0) + \epsilon\rho(t_0)x'^a(t_0) = 0.\end{gathered}$$

Consider first the case $\gamma\delta = 0$, in which we have $\epsilon \neq 0$, by § 9. When the point x is not at a corner, the situation may always be reduced to this case. Then obviously Θ must vanish for $t = \tau$, τ^-, or τ^+. In case x is at a corner and $\gamma\delta > 0$, $\rho(\tau) = 0$, we must have

$$\begin{vmatrix} \xi_i{}^a(\tau) & \eta_i{}^a(\tau) & [\gamma x'^a(\tau^-) + \delta x'^a(\tau^+)] & 0 \\ \xi_i{}^a(t_0) & \eta_i{}^a(t_0) & 0 & x'^a(t_0) \end{vmatrix} = 0,$$

* *Loc. cit.*, p. 206.

P14765

from which we see that Θ changes sign for $t = \tau$.

To prove the converse, suppose, in the first place, that Θ vanishes for $t = \tau$. Then the equations

$$\begin{aligned} a^i\xi_i{}^a(\tau) + b^i\eta_i{}^a(\tau) + \gamma x'^a(\tau) \qquad\qquad &= 0 \\ a^i\xi_i{}^a(t_0) + b^i\eta_i{}^a(t_0) \qquad\qquad + \delta x'^a(t_0) &= 0 \end{aligned}$$

have solutions a^i, b^i, γ, δ, with the constants a^i, b^i, not all zero. Let $\xi^a = a^i\xi_i{}^a + b^i\eta_i{}^a + \rho x'^a$, where $\rho(t_0) = \delta$, $\rho(\tau) = \rho(t_1) = \cdots = \rho(t_m) = 0$. Then $\xi^a(t_0) = 0$, $\xi^a(\tau) = -\gamma x'^a(\tau)$. By lemmas 10.4 and 10.2, ξ^a is not of the form $\rho x'^a$. Consequently the point x corresponding to $t = \tau$ is conjugate to x_0. Suppose finally that Θ changes sign without vanishing at $t = \tau$. Then $\gamma\Theta(\tau^-, t_0) + \delta\Theta(\tau^+, t_0) = 0$, where $\gamma\delta > 0$. Hence the equations

$$\begin{aligned} a^i\xi_i{}^a(\tau) + b^i\eta_i{}^a(\tau) + \gamma_1[\gamma x'^a(\tau^-) + \delta x'^a(\tau^+)] &= 0 \\ a^i\xi_i{}^a(t_0) + b^i\eta_i{}^a(t_0) \qquad\qquad + \delta_1 x'^a(t_0) &= 0 \end{aligned}$$

have solutions a^i, b^i, γ_1, δ_1, not all zero. The constant γ_1 is not zero, since otherwise the determinant Θ would vanish at τ. If the constants a^i, b^i, are all zero, we have $\gamma x'^a(\tau^-) + \delta x'^a(\tau^+) = 0$, and the solution ξ^a of the Jacobi equations occurring in the definition of conjugate point may be taken arbitrarily. If the constants a^i, b^i, are not all zero, let $\xi^a = a^i\xi_i{}^a + b^i\eta_i{}^a + \rho x'^a$, where $\rho(t_0) = \delta_1$, $\rho(t_1) = \cdots = \rho(t_m) = 0$. Then ξ^a is a solution of the Jacobi equations, not of the form $\rho x'^a$, having

$$\xi^a(t_0) = 0, \qquad \xi^a(\tau) = -\gamma_1\gamma x'^a(\tau^-) - \gamma_1\delta x'^a(\tau^+), \qquad \gamma_1{}^2\gamma\delta > 0.$$

Suppose we have a one-parameter family of extremaloids $x^a = \phi^a(t, c)$ containing the extremaloid E_0 for $c = c_0$, and with the corners all occurring at the fixed parameter values $t_1, \cdots, t_m$. If the functions ϕ^a are substituted for x^a in the Euler equations (2.1) and the corner conditions (2.2), these equations become identities in t and c. If we differentiate these identities with respect to c and set $c = c_0$, we find that $\xi^a = (\partial/\partial c)\phi^a(t, c_0)$ is a solution of the Jacobi equations (8.1) for the extremaloid E_0. Hence from the $2n-2$ parameter family of extremaloids $x^a = \Phi^a(t, a, b)$ discussed in § 3 (supposing that the corners all occur at the fixed parameter values t_k), we obtain a set of $2n-2$ solutions of the Jacobi equations

$$\xi_i{}^a(t) = (\partial/\partial a^i)\Phi^a(t, a_0, b_0), \qquad \eta_i{}^a(t) = (\partial/\partial b^i)\Phi^a(t, a_0, b_0).$$

For this set the determinant $\mathfrak{W}$ does not vanish on at least one extremal arc of E_0, by the imbedding theorems for extremals, and hence $\mathfrak{W}$ does not vanish on E_0, by lemma 10.1. The corresponding determinant Θ does not vanish

identically, by lemma 10.4. As explained in the closing paragraph of the sign of each of the determinants

$$(10.1)\qquad \begin{vmatrix} \dfrac{\partial\Phi^\alpha(t)}{\partial a^i} & \dfrac{\partial\Phi^\alpha(t)}{\partial b^i} & \dfrac{\partial\Phi^\alpha(t)}{\partial t} & 0 \\ \dfrac{\partial\Phi^\alpha(t_0)}{\partial a^i} & \dfrac{\partial\Phi^\alpha(t_0)}{\partial b^i} & 0 & \dfrac{\partial\Phi^\alpha(t_0)}{\partial t} \end{vmatrix} \begin{matrix} a \\ b \end{matrix}$$

$$(10.2)\qquad \begin{vmatrix} \dfrac{\partial\phi^\alpha}{\partial a^i} & \dfrac{\partial\phi^\alpha}{\partial t} \end{vmatrix}_{a = a_0},$$

is unaffected by the transformations necessary to make the corners ap... fixed parameter values. If the parameters a and b are properly sele...d n 1 parameter family of extremaloids $x^\alpha = \phi^\alpha(t, a)$ passing thro... fixed point x_0 may be obtained from the general family by setting b i. e., $\phi^\alpha(t, a) = \Phi^\alpha(t, a, b_0)$. Hence the determinant (10.1) is a multip... the determinant (10.2) with multiplier not zero. Lemma 10.5 comb... with these facts yields the

THEOREM. *Let $x^\alpha = \Phi^\alpha(t, a, b)$ be equations of the general $2n$ - 2 ... meter family of extremaloids, and $x^\alpha = \phi^\alpha(t, a)$ be equations of the ... parameter family passing through the point x_0, in which the extrem...l is contained for $a = a_0$, $b = b_0$. Then the points x conjugate to x_0 ... correspond to the values of the parameter t for which the determinants (1... and* (10.2) *vanish or change sign.*

SUFFICIENT CONDITIONS FOR A MINIMUM.

11. *Definition of a field, and a general sufficiency theorem.** A ... *with m discontinuities* is defined to be a continuum $\mathfrak{F}$ in x-space, ha... associated with it m corner-manifolds

$$S_k : x^\alpha = X_k{}^\alpha(a)$$

of $n - 1$ dimensions, and n slope-functions $p^\alpha(x)$, with the properties:

1) the corner manifolds S_k are non-singular, have no multiple p... do not intersect each other, and each S_k divides $\mathfrak{F}$ into two parts:

2) the slope functions $p^\alpha(x)$ do not vanish simultaneously, a d their first derivatives are continuous between corner manifolds S_k an... proach finite limits on each side of each S_k;

* Cf. Bliss, *The Problem of Lagrange in the Calculus of Variations*, mimeogr... lectures prepared by O. E. Brown, Chap. IV. Also Bliss, "The Transformati... Clebsch in the Calculus of Variations," *Proceedings of the Toronto Congress*, p. ...

3) the two limits p^{a-} and p^{a+} of the functions p^a at a point of a manifold S_k always determine directions on the same side of S_k and never tangent to it;

4) the functions $\partial F[x, p(x)]/\partial x'^a$ are continuous in $\mathfrak{F}$;

5) the integral

$$J^* = \int \frac{\partial F[x, p(x)]}{\partial x'^a} \, dx^a$$

is independent of the path in $\mathfrak{F}$.

The condition 3) is equivalent to saying that the two determinants

$$| \partial X_k^a/\partial a^i \quad p^{a-} |, \qquad | \partial X_k^a/\partial a^i \quad p^{a+} |,$$

are both positive or both negative along a corner manifold. It may be proved in the usual way that the solutions of the equations

$$dx^a/dt = p^a(x)$$

between manifolds S_k are all extremals. The condition 4) implies that these extremal arcs may be pieced together to form extremaloids. These are called the *extremaloids of the field.* Along an extremaloid of the field the integrals J and J^* obviously have the same value. From these facts we may readily prove the

GENERAL SUFFICIENCY THEOREM. *Let E be an extremaloid of a field $\mathfrak{F}$ in which the Weierstrassian function*

$$\mathcal{E}[x, p(x), q] \geqq 0$$

for all directions q. Then if C is an admissible curve lying in $\mathfrak{F}$ and joining the same two points as E, we must have

$$J(C) \geqq J(E).$$

If, except for points x on the manifolds S_k, the function $\mathcal{E}$ vanishes only when the direction q coincides with the direction $p(x)$ (i. e., *if the extremaloids of the field are positively strong*), *then $J(C) = J(E)$ only when the curve C coincides with the extremaloid E.*

This theorem is applicable even in cases where the hypotheses $F_1 \neq 0$, $\Omega_0 \neq 0$, used in deriving necessary conditions, are not everywhere fulfilled.

12. *Sufficient conditions for the existence of a field.** The theorem we wish to secure is the following:

* Bolza gives a geometric proof for the plane case in *Vorlesungen über Variationsrechnung*, pp. 381 ff.

Let E_0 be an extremaloid along which $F_1 \neq 0$, and $\Omega_0 \neq 0$ *at the co...
. f E_0 contains no multiple point and no point conjugate to one of its e...
points, then E_0 is an extremaloid of a field $\mathfrak{F}$, as defined in* § 11.

From the discussion of § 3 and § 10 we know that E_0 is a member ... $2n-2$ parameter family of extremaloids

$$x^a = \Phi^a(t, a, b),$$

for which the determinant

$$\Theta(t, t_0) = \begin{vmatrix} \dfrac{\partial \Phi^a(t)}{\partial a^i} & \dfrac{\partial \Phi^a(t)}{\partial b^i} & \dfrac{\partial \Phi^a(t)}{\partial t} & 0 \\ \dfrac{\partial \Phi^a(t_0)}{\partial a^i} & \dfrac{\partial \Phi^a(t_0)}{\partial b^i} & 0 & \dfrac{\partial \Phi^a(t_0)}{\partial t} \end{vmatrix}$$

does not vanish nor change sign for $t_0 < t \leqq T$. By lemma 10. 3 we h... that for t and $\bar{t}_0$ in a sufficiently small neighborhood $t_0 - \gamma \leqq t \leqq t_0 + \gamma$ of ...

$$(12.1) \qquad \Theta(t, \bar{t}_0) = (t - \bar{t}_0)^{n-1} \lambda(t, \bar{t}_0), \qquad \lambda(t, \bar{t}_0) \neq 0.$$

The function $\Theta(t, \bar{t}_0)$ is continuous for $t_0 - \gamma \leqq \bar{t}_0 \leqq t_0 + \gamma$, and fo... each of the intervals $t_0 + \gamma \leqq t \leqq t_1$, $t_k \leqq t \leqq t_{k+1}$, ($k = 1, \cdots, m$; $t_{...}$ Since Θ preserves its sign for this range of t and $\bar{t}_0 = t_0$, Θ will still p... its sign for a fixed value $\bar{t}_0$ such that $t_0 - \gamma < \bar{t}_0 < t_0$, and for the same ... of t. The combination of this result with (12. 1) shows then that $\Theta(t$... does not vanish nor change sign for $\bar{t}_0 \leqq t \leqq T$.

Let $\bar{x}_0$ be the point on E_0 (preceding x_0) corresponding to $\bar{t}_0$, and let

$$x^a = \phi^a(t, a)$$

be equations of the $n-1$ parameter family of extremaloids passing thro... $\bar{x}_0$ for $t = \bar{t}_0$, containing E_0 for $a = a_0$. We now suppose also that the pa... meter t is chosen so that the corners on all the extremaloids occur for fixed parameter values $t_1, \cdots, t_m$. From the relation between the de... minants $\Theta(t, \bar{t}_0)$ and

$$D(t, a) = | \, \partial \phi^a / \partial a^i \qquad \partial \phi^a / \partial t \, |$$

discussed in § 10, we know that $D(t, a_0)$ does not vanish nor change sign ... $t_0 \leqq t \leqq T$.

The next step is to show that these facts imply the existence of a un... solution $t = t(x)$, $a^i = a^i(x)$, of the equations $x^a = \phi^a(t, a)$, for poin... in a neighborhood of the extremaloid E_0, so that there is a contin... containing the extremaloid E_0 which is in one-to-one continuous corresp... ence with a neighborhood of the set of points $t_0 \leqq t \leqq T$, $a^i = a_0^i$.

ordinary implicit function theorems are not immediately applicable, since the derivatives of the functions ϕ^a are not all continuous. Let us recall, however, from the method of constructing the family of extremaloids, that this family may be broken up into $m + 1$ families of extremals

$$x^a = \psi_k{}^a(t, a) = \phi^a(t, a) \qquad (t_{k-1} \leqq t \leqq t_k).$$

Moreover, these extremals may be slightly extended beyond the corner manifolds so that the functions $\psi_k{}^a$ have first and second partial derivatives which are continuous, and the determinants

$$D_k(t, a) = | \, \partial\psi_k{}^a/\partial a^i \qquad \partial\psi_k{}^a/\partial t \, |$$

do not vanish and all have the same sign. The standard implicit function theorems are therefore applicable to each of these families. For points x near a corner x_k, there will be a unique solution $t^-(x)$, $a^{i-}(x)$ near t_k, $a_0{}^i$, of the equations $x^a = \psi_k{}^a(t, a)$, and a unique solution $t^+(x)$, $a^{i+}(x)$, of the equations $x^a = \psi^a{}_{k+1}(t, a)$. If either t^- or t^+ equals t_k, then $t^- = t^+ = t_k$, $a^{i-} = a^{i+}$, since

(12.2) $$\psi_k{}^a(t_k, a) = \psi^a{}_{k+1}(t_k, a).$$

The remaining cases are:

1) $t^- < t_k$, $t^+ < t_k$;
2) $t^- > t_k$, $t^+ > t_k$;
3) $t^- < t_k$, $t^+ > t_k$;
4) $t^- > t_k$, $t^+ < t_k$.

We shall show that in a sufficiently small neighborhood of (x_k, t_k, a_0) the third and fourth cases cannot occur, so that the equations $x^a = \phi^a(t, a)$ define functions $t(x)$, $a^i(x)$, which are single-valued near x_k. In the determinants D_k, D_{k+1}, suppose the arguments (t, a) of each element to be replaced by independent variables, and denote the resulting functions of n^3 variables by $\mathfrak{D}_k$, $\mathfrak{D}_{k+1}$. Since $\mathfrak{D}_k = D_k(t_k, a_0)$, $\mathfrak{D}_{k+1} = D_{k+1}(t_k, a_0)$, when the arguments of all the elements are set equal to (t_k, a_0), there will be a neighborhood of (t_k, a_0) in which $\mathfrak{D}_k$ and $\mathfrak{D}_{k+1}$ do not vanish and have the same sign. Suppose that either the third or the fourth case listed above occurs in this neighborhood. Then by using equations (12.2) we have

(12.3) $$\begin{aligned} 0 &= \psi^a{}_{k+1}(t^+, a^+) - \psi_k{}^a(t^-, a^-) \\ &= \psi^a{}_{k+1}(t^+, a^+) - \psi^a{}_{k+1}(t_k, a^+) \end{aligned}$$

$$
\begin{aligned}
&+ \psi_k^a(t_k, a^-) - \psi_k^a(t^-, a^-) \\
&+ \psi_k^a(t_k, a^+) - \psi_k^a(t_k, a^-) \\
= &\left[\frac{\partial \psi^a_{k+1}(\bar{t}^+, a^+)}{\partial t} \left(\frac{t^+ - t_k}{t^+ - t^-}\right) + \frac{\partial \psi_k^a(\bar{t}^-, a^-)}{\partial t} \left(\frac{t_k - t^-}{t^+ - t^-}\right)\right] (t^+ - t^-) \\
&+ \frac{\partial \psi_k^a(t_k, \bar{a})}{\partial a^i} (a^{i+} - a^{i-}),
\end{aligned}
$$

where the arguments $\bar{t}^+$, $\bar{t}^-$, $\bar{a}$, are those prescribed by the theorem of the mean From (12. 2) we have also

$$
\frac{\partial \psi_k^a(t_k, a)}{\partial a^i} = \frac{\partial \psi^a_{k+1}(t_k, a)}{\partial a^i},
$$

so that the determinant of the matrix of coefficients of the quantities $(t^+ - t^-)$, $(a^{i+} - a^{i-})$ in equations (12. 3) reduces to

$$
\left(\frac{t^+ - t_k}{t^+ - t^-}\right) \mathfrak{D}_{k+1} + \left(\frac{t_k - t^-}{t^+ - t^-}\right) \mathfrak{D}_k \neq 0.
$$

Hence the third and fourth cases cannot occur. The extended implicit function theorem of Bolza can now be proved for the present situation by exactly the same arguments as those used by Bliss for the case of continuous derivatives.*

We can now define the slope functions for the field by the equations

$$
p^a(x) = (\partial/\partial t)\phi^a[t(x), a(x)].
$$

The five properties in the definition of a field in § 11 are readily verified. The third follows from the fact that the determinant $D(t, a)$ does not vanish nor change sign in a sufficiently restricted neighborhood of the extremaloid E_0. To verify property 5), the invariance of the Hilbert integral, we compute the partial derivatives of the *field-integral*

$$
W(x) = \int_{\bar{t}_0}^{t(x)} F\{\phi[t, a(x)],\ (\partial/\partial t)\phi[t, a(x)]\}dt
$$

and find that between corner manifolds they are

$$
\partial W/\partial x^a = (\partial/\partial x'^a)F[x, p(x)].
$$

* Cf. *Princeton Colloquium Lectures*, pp. 20, 21.

Since these derivatives are continuous even on the corner manifolds, we see at once that the Hilbert integral J^* is independent of the path, for all admissible curves in $\mathfrak{F}$.

13. *A set of sufficient conditions for a relative minimum.* From the preceding sections we can now readily prove that the following conditions are sufficient for $J(E_0)$ to be a *proper strong relative* minimum.*

I. *E_0 is an extremaloid having no multiple point.*

II. $\Omega_0 \neq 0$ *at the corners on E_0.*

III. *E_0 is positively strong, and is uniquely determined by each of its elements (x, x').*†

IV. *E_0 contains no point conjugate to one of its end-points.*

For, the third condition implies $F_1 \neq 0$ along E_0, and this with the other conditions implies that E_0 is an extremaloid of a field, by § 12. Then we apply lemmas 5. 3 and 5. 6 to show that the extremaloids of the field are all positively strong if the field is restricted to a sufficiently small neighborhood of E_0. This enables us to apply the general sufficiency theorem of § 11.

The third condition is understood to mean that the function $\mathcal{E}_1$ of § 5 does not vanish at either of the end-points x_0, X, of the extremaloid E_0.

* For the terminology, cf. Bolza, *loc. cit.*, pp. 17, 197.

† See § 5, p. 11.

Geodesics on a Toroid.*

By B. F. Kimball.

1. *Introduction.* The study of the geodesics on a toroid is of interest not only because of the geometric properties of the surface which this study discloses, but also because the geodesic curves are the solutions of a differential equation which is of interest; namely,

$$v'(u) = [r(v)/\alpha]\,[r^2(v) - \alpha^2]^{1/2}$$

where α is a constant not zero and $r(v)$ is a real, analytic, periodic function of v greater than zero. The function $r(v)$ is restricted to have only one maximum and one minimum value in a period interval. Taking $v = 0$ at a maximum of r it is also required that $r(v) = r(-v)$.

The results obtained in this paper through Part III apply to solutions of the above differential equation, the solutions being mapped on a surface of revolution. The results set forth in Part IV apply to a slightly more restricted differential equation [see condition (c) § 2]. The reason for introducing such further restriction is that a more complete and interesting study of the envelopes of the geodesics on the surface is thereby made possible.

The surface of revolution defined in § 2 is called a toroid since it resembles a torus. The torus was defined by Cayley † as "a surface generated by the rotation of a conic about a fixed axis anywise situate." The geodesics on the anchor ring torus have already been studied by G. A. Bliss.‡

The differential equation of the geodesics on the anchor ring, Bliss solved by means of elliptic functions and in an elegant manner studied the solutions thus obtained. A destinctive feature of the present paper is the policy of dealing directly with the differential equations, since closed solutions are not obtainable in this more general case, and of getting at the desired properties of the solutions directly from them. This paper uses for the first time, as far as the writer is aware, the equation of second variation as an aid in a specific investigation of the envelopes of geodesics, and makes application of the envelope theory of Lindeberg (cf. § 16) to that study. This introduction

* Presented to the American Mathematical Society, October 30, 1926.

† Cayley, *Mathematical Papers*, Vol. 7, p. 246, and Vol. 8.

‡ "The Geodesic Lines on the Anchor Ring," *Annals of Mathematics*, Ser. 2 Vol. 4 (1903), p. 1.

would not be complete without making an acknowledgment of the inspiration gained from the work of Prof. Marston Morse in this field of study.

PART I. THE SURFACE.

2. *Definition of the surface.*

(a) *The surface is taken as a surface of revolution with the z-axis as its axis of revolution. The generatrix is a simple, closed curve which lies initially in the xz-plane and does not cut the z-axis. It will be represented in terms of two functions $r(v)$ and $z(v)$ which are real, analytic, periodic functions of v for all real values of v and with a period 2ω. In its initial position it is expressed in the form*

$$(2.1) \qquad \begin{aligned} x &= r(v) \qquad & r(-\omega) &= r(+\omega), \\ z &= z(v) \qquad & z(-\omega) &= z(+\omega), \end{aligned}$$

where v is the arc length of the curve. Only values of v in the interval $-\omega \leqq v \leqq +\omega$ will be considered.

(b) *It shall further be required that the curve (2.1) be symmetrical with respect to the x-axis.* Of the intersections of the generating curve with the x-axis, let A be the one of maximum x and B the one of minimum x. The points A and B divide the generatrix into two segments. Of these two segments at least one has some points at which z is positive. Let C be a point on this segment at which z takes on its absolute maximum. Let D be the reflection of C in the x-axis. The arc length v used in (2.1) will be measured in the sense $A\,C\,B$ from A as origin.

(c) The last important restriction on the above surface is that *the total curvature K of the surface satisfy the relation $dK/dv < 0$ at all points of $A\,C\,B$ except at A* and *B.*

3. *Some properties of the surface.* One can write the equations of the surface in the form:

$$x = r(v) \cos u, \quad y = r(v) \sin u, \quad z = z(v),$$

where u is the angle that the half meridian plane through the point makes with the xz-plane measured in the xy-plane from the x-axis in a sense which is counterclockwise from the point of view of a position on the positive *z-axis.* Then the square of the element of arc on the surface will be

$$(3.1) \qquad ds^2 = dv^2 + r^2(v)\,du^2.$$

It follows from a familiar formula of Differential Geometry * that

* Eisenhart, *Differential Geometry* (1909), p. 208, VI (55).

$$K = -r''(v)/r(v). \tag{3.2}$$

Let k be the curvature of the generating curve taken with the sign given by the curvature formula:

$$k = r'z'' - z'r''. \tag{3.3}$$

Also since v is the arc length

$$r'^2 + z'^2 = 1. \tag{3.4}$$

This gives

$$r'r'' + z'z'' = 0. \tag{3.5}$$

Multiplying (3. 3) by $-z'$, and substituting for $z'z''$ its value obtained from (3. 5) one obtains

$$-z'k = r'' \tag{3.6}$$

and (3.2) becomes

$$K = z'k/r. \tag{3.7}$$

From hypothesis (a), (b) and (c) and the above formulae one obtains the following theorem concerning the properties of the generating curve. The will follow.

THEOREM 1. (a) *The curvature k of the generatrix is always positive. The derivative $r'(v)$ vanishes only twice on the generatrix; namely, at points A and B and at these points r takes on its absolute maximum and absolute minimum respectively.*

Proof. The proof of statement (a) follows from the consideration of formula (3.7). Since $dK/dv < 0$ between A and B one sees that K cannot vanish more than once on $A\,C\,B$ (including the end points A and B). Hence from the above formula z' cannot vanish more than once. But z' necessarily vanishes at the maximum C. Thus K is zero on $A\,C\,B$ only at the point C. It then follows that if k vanishes at all on $A\,C\,B$, it can only vanish at C. One can now show that k cannot vanish at this point. Differentiating (3.7) and employing hypothesis (c) one obtains the relation

$$\frac{dK}{dv} = \frac{r(z'k' + z''k) - z'kr'}{r^2} < 0, \tag{3.8}$$

which shows that k and z' cannot vanish at the same point. Thus k does not vanish at C and from symmetry cannot vanish at any point on the segment $A\,D\,B$. Thus statement (a) is proved. To determine the sign of k observe that, at the point A,

$$z' = 1, \quad r' = 0, \quad r'' < 0.$$

Formula (3.3) then shows that k is positive at A and hence positive at all points of the curve.

(b) The statement (b) can be proved as follows. Consider the generatrix in the xz-plane. Let θ denote the angle which the tangent directed in the sense of increasing v makes with the positive x-axis. For a simple closed curve possessing a continuously turning tangent at every point it has been proved * that this angle changes by a net amount of 2π as the curve is traversed once. Now for above curve $k = d\theta/dv > 0$. Hence θ must increase monotonically, and by an amount 2π as the curve is traversed from D to D in the positive sense. Thus $\cos\theta$ is zero only twice; namely, at A and B. The statement (b) above follows.

If the generatrix is an ellipse with axis $A\,B$ equal to $2a$ and axis $C\,D$ equal to $2b$ and r_0, the distance from the center of the ellipse to the z-axis, greater than a; it can be shown that necessary and sufficient conditions that $dK/dv < 0$ for the interval $0 < v < \omega$ are that

$$a \geqq b,$$

or

$$a < b < (2/3^{1/2})a, \quad r_0 \geqq 4a(b^2 - a^2)/(4a^2 - 3b^2). \tag{3.9}$$

PART II. THE GEODESICS.

4. *The differential equations of the geodesics.* It was seen that for the system of coordinates chosen, the square of the element of arc on the surface was

$$ds^2 = dv^2 + r^2(v)\,du^2. \tag{4.1}$$

The differential equation of the geodesics in parametric form is thus †

$$(\dot u\,\ddot v = \ddot u\,\dot v) - (2\,r'/r)\dot u\,\dot v^2 - r\,r'\,\dot u^3 = 0 \tag{4.2}$$

where a dot above a letter indicates differentiation with respect to the parameter t and r' is dr/dv. One also needs the equation of the geodesics in non-parametric form for the case where u is taken as the independent variable. Setting $t = u$, equation (4.2) becomes

* Watson, *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 15 (1916), pp. 227-42.

† Bolza, *Variationsrechnung* (1909), § 26, p. 210. Future references to this book will be indicated by the letter B.

(4.3) $$v''(u) - [2r'(v)/r(v)]\,v'^2(u) - r(v)r'(v) = 0.$$

Other formulae for the geodesics are found by integrating one of the Euler equations giving

(4.4) $$r^2\dot{u}/[\dot{v}^2 + r^2(v)\dot{u}^2]^{1/2} = \alpha,$$

where α is a constant of integration. Write (4.4) in the form

(4.5) $$r^2 du/ds = \alpha$$

or

(4.6) $$r\cos\chi = \alpha$$

where χ is that angle which the positive direction of a geodesic makes with the positive tangent to the parallel thru the point on the geodesic considered, measured in a counterclockwise sense from the point of view of a position on the exterior normal to the surface. For geodesics along which u may be taken as the independent variable, when one sets $t = u$ (4.4) becomes

(4.7) $$r^2/(v'^2 + r^2)^{1/2} = \alpha,$$

which solved for v' gives

(4.8) $$v' = \pm (r/\alpha)(r^2 - \alpha^2)^{1/2}.$$

Now the equation of a geodesic must be a solution of (4.4) and therefore of equations (4.5) — (4.8) (cf. B. § 26); but all curves that are solutions of these equations are not geodesics—*e. g.*, any parallel will be a solution of (4.6), while in order that it be a solution of the general equation (4.2) it is necessary that r' be zero. Incidentally it is seen that *the only parallels that are geodesics are the inside and outside equators, that is, the curves* $v = 0$ *and* $v = \omega$.

5. *The symmetry of the geodesics.* Taking note of the fact that $r(v)$ is an even function of v, it is seen from the form of the equation (4.2) that

(5.1) *the system of geodesics is symmetrical with respect to the equatorial plane of the surface;*

(5.2) *it is symmetrical with respect to any meridian plane;*

(5.3) *if the surface be rotated about an axis of revolution through any angle, each geodesic on the surface in its original position will coincide at the end of the rotation with a geodesic on the surface in its final position.*

6. *The classification of the geodesics.* Consider the geodesics issuing from a given point (u_0, v_0) at an angle χ restricted to the values

(6.1) $$-\pi/2 < \chi \leqq +\pi/2.$$

Clearly a reflection of these geodesics in the meridian plane through (u_0, v_0) will give all the other geodesics through (u_0, v_0). Let the maximum and minimum values of r on the generatrix be respectively r_2 and r_1. The geodesics issuing from (u_0, v_0) with the angles χ given by (6.1) will determine values of α in (4.6) such that $0 \leqq \alpha \leqq r_2$. These geodesics will be classified as follows:

Class I	$\alpha = 0,$
Class II	$0 < \alpha < r_1,$
Class III	$\alpha = r_1,$
Class IV	$r_1 < \alpha < r_2,$
Class V	$\alpha = r_2.$

7. *The nature of the geodesics.* The following theorems may be verified by using (4.6), (4.8) and the well-known fact that there is a unique geodesic through a given point in a given direction on the surface. In the case of Theorem 5 formula (4.3) will also be needed. Since the results are not essentially new the proofs of these theorems are omitted.*

THEOREM 2. *The meridians ($u =$ const.) are the only geodesics which belong to Class I.*

THEOREM 3. *On a geodesic of Class II, v must increase or decrease without limit as u increases without limit, and the geodesic will cross both inside and outside equators an infinite number of times.*

THEOREM 4. *Geodesics of Class III consist of (A) the inside equator, and (B) two geodesics through every point of the surface which approach the inside equator asymptotically in either sense but never cross it. The cosine of the angle χ_1 at which a geodesic of Class III (B) crosses the outside equator is given by the formula* $\cos \chi_1 = r_1/r_2$.

THEOREM 5.† *Each geodesic of Class IV lies in the region on the surface for which $r \geqq \alpha$ and becomes successively tangent to the two parallels $r = \alpha$ as u increases.*

THEOREM 6. *The only geodesic in Class V is the outside equator.*

---

* Forsythe, *Differential Geometry* (1912), § 93, p. 132.

† For a further description of geodesics of Class IV see Theorem 14.

PART III. THE CONJUGATE POINTS.

8. *The formulae needed for the discussion of the conjugate points.* There will be occasion to speak of the rth conjugate point to a point P on a given geodesic E. Let the coordinates of this point be (u_0, v_0). Consider a solution $\bar{w}(u)$ of the Jacobi differential equation (B. § 29) set up for the geodesic E such that

$$\bar{w}(u_0) = 0, \qquad \bar{w}'(u_0) = 1.$$

From the theory of the second order differential equation* it is known that the zeros of $\bar{w}(u)$ are discrete. The first conjugate point to P on E is the zero of $\bar{w}(u)$ corresponding to the smallest value of u greater than u_0, the second conjugate point to P on E is the second zero of $\bar{w}(u)$ for u greater than u_0, and in general the rth conjugate point to P is the rth zero of $\bar{w}(u)$ on E for u greater than u_0. The words "conjugate point" will be used in this article to designate the first conjugate point as defined above. Conjugate points denoted by $P_1, P_2, \cdots, P_r$, will refer to the first, second, $\cdots$, rth conjugate points according as the subscripts are $1, 2, \cdots, r$.

It will be convenient to make use of the following facts derivable from the general theory of the Calculus of Variations. If there be given any geodesic segment G without double points on the surface, it is possible (B. § 40) to choose a set of surface coordinates (s, n) as follows. The curves $s = \text{const.}$ may be taken as geodesics perpendicular to G, and n can be taken as the distance measured along these geodesics from G in one of the two possible senses chosen arbitrarily, while s is the distance measured along G from a given point in a given sense. The domain in which it is possible to take such a set of coordinates and still retain the one-to-one correspondence between the pairs (s, n) and the points on the surface will in general have to be limited to a sufficiently small region R neighboring G, but including G entirely within the interior. For the coordinates (s, n) the first fundamental quadratic form of the surface will be

$$d\sigma^2 = dn^2 + C^2(s, n)\,ds^2 \tag{8.1}$$

(B. § 28) where $C(s, n)$ is a real positive analytic function of s and n in R and

$$C(s, 0) = 1, \quad C_n(s, 0) = 0, \quad C_s(s, 0) = 0. \tag{8.2}$$

Under such a choice of a coordinate system consider an analytic family of geodesics

$$n = \Phi(s, \alpha), \quad 0 = \Phi(s, 0), \tag{8.3}$$

* Bôcher, *Leçons sur les Méthodes de Sturm*, § 12, p. 43.

where Φ is analytic in s and α in the neighborhood of a segment G of the geodesic $\alpha = 0$. Let $K(s, n)$ denote the total curvature of the surface as a function of s and n. The Jacobi differential equation, otherwise known as the equation of first variation,* for this geodesic G will then take the form

$$w''(s) + K(s, 0) w(s) = 0. \tag{8.4}$$

It can be shown that the equation of the second variation (B. § 8b) will be

$$\Phi''_{\alpha\alpha}(s, 0) + K(s, 0)\Phi_{\alpha\alpha}(s, 0) = -K_n(s, 0)\Phi_\alpha^2(s, 0), \tag{8.5}$$

and both of these equations will hold at all points of the segment G.

9. *Conjugate points on geodesics of Class I.* The only geodesics of this class are the meridian ovals. *No point on a meridian has a conjugate point.*

Along any meridian one can regard v as the independent variable. The complete set of meridian ovals is a family of geodesics given by the equation

$$u = \Phi(v, \beta)$$

where $\Phi(v, \beta)$ has the simple form

$$\Phi(v, \beta) = \beta,$$

and β is an arbitrary constant. The Jacobi equation in the nonparametric form with v the independent variable and w the dependent variable can be set up for any particular meridian and according to the general theory it will have for one particular solution

$$w = \Phi_\beta(v, \beta) = 1.$$

Thus this particular solution never vanishes. Therefore by Sturm's separation theorem (B. § 11) the solution of the Jacobi equation which determines the conjugate point P_1 to a given point P on the meridian considered cannot vanish twice. Hence no such conjugate point to P exists on a meridian. It may be noted that the above proof would apply to any analytic, regular surface of revolution for which the generatrix is a closed curve.

10. *Conjugate points on geodesics of Class II. There are no conjugate points on geodesics of this class.*

Any geodesic of this class can be represented in the form $v = v(u)$, where $v(u)$ is an analytic function of u for all real values of u. Symmetry considerations show that the geodesics obtained by rotating the given geo-

---

* Darboux, *Théorie des Surfaces*, Vol. 3, § 627, p. 97.

desic through any angle about the axis of revolution of the surface will also be geodesics of Class II. The family of geodesics so obtained can be represented as follows:

$$v = \Phi(u, \beta) = v(u + \beta),$$

where β is an arbitrary constant. The Jacobi equation of the nonparametric case set up for any one of these geodesics will have as a particular solution

$$\Phi_\beta = v'(u + \beta).$$

According to formula (4.8) this derivative is never zero for a geodesic of Class II. Again applying Sturm's separation theorem one obtains the result to be proved.

11. *Conjugate points on geodesics of Class III.*

(A) *There are no conjugate points corresponding to points on the inside equator.* Reference to formula (3.7) shows that the surface curvature is a negative constant on the inside equator. Accordingly a solution of the Jacobi equation (8.4) which vanishes once and is not identically zero never vanishes again. Thus there are no conjugate points corresponding to points on the inside equator.

(B) Geodesics of Class III (B) are geodesics asymptotic to the inside equator. *These geodesics have no conjugate points.* This can be proved by a method similar to that used in the proof of the non-existence of conjugate points on geodesics of Class II. Thus the theorem:

THEOREM 7. *No point on a geodesic of Class I, II, or III has a conjugate point.*

12. *Conjugate points on geodesics of Class IV. On a geodesic of Class IV the rth extremum of v following a given extremum of v is the rth conjugate point to that extremum.* To prove this statement one proceeds as in the discussion of geodesics of Classes II and III(B). If the given geodesic is represented in the form $v = v(u)$, then

$$v = v(u + \beta)$$

will represent a family of geodesics of this class, while $v'(u)$ will be a solution of the Jacobi equation corresponding to the given geodesic. This solution $v'(u)$ vanishes at the extrema of $v(u)$ and only at these extrema, and the statement is proved.

The following lemma of the Calculus of Variations will be needed:

LEMMA 1. *If Q_1, Q_r and Q_{r+1} are respectively the first, rth and $r+1$th conjugate points (cf. § 8) to the point Q on a given geodesic E corresponding to a regular problem in the Calculus of Variations, then as a point P varies on E from Q to Q_1 moving always in the same sense, its rth conjugate point P_r will move continuously from Q_r to Q_{r+1}, moving always in the same sense.*

In order to prove this lemma, take s as the arc length measured along the given geodesic in the sense that leads from Q to Q_1. Let $w_1(s)$ and $w_2(s)$ be any two independent solutions of the Jacobi equation. Let s and s_r denote the values of s at P and P_r and denote by a and a_r the values of s at Q and Q_r respectively. Then

$$(12.1) \qquad \begin{vmatrix} w_1(s) & w_2(s) \\ w_1(s_r) & w_2(s_r) \end{vmatrix} = 0,$$

and if one shows that equation (12.1) can be solved for s_r as a continuous, monotonically increasing function of s, for s in the interval

$$(12.2) \qquad a \leqq s \leqq a_1$$

corresponding to the extremal from Q to Q_1, the lemma will be demonstrated. Regard the determinant (12.1) as a function $\phi(s, s_r)$. Consider a particular pair of the conjugate points $(s^0, s_r{}^0)$ which is thus a solution of $\phi(s, s_r) = 0$, where s^0 is a value of s in the interval (12.2). In order to show that

$$(12.2) \qquad \phi_{s_r}(s^0, s_r{}^0) = \begin{vmatrix} w_1(s^0) & w_2(s^0) \\ w'_1(s_r{}^0) & w'(s_r{}^0) \end{vmatrix} \neq 0,$$

observe that

$$\begin{vmatrix} w_1(s^0) & w_2(s^0) \\ w_1(s) & w_2(s) \end{vmatrix}$$

may be regarded as a function $u(s)$ which with s^0 constant is a solution of the Jacobi equation that vanishes for $s = s_r{}^0$. Moreover $u(s)$ is not identically zero, for if it were, $w_1(s)$ and $w_2(s)$ would be linearly dependent. Furthermore $u'(s_r{}^0)$ cannot then be zero since $u(s_r{}^0) = 0$. But $u'(s_r{}^0)$ is the determinant in (12.3). It is thus proved not zero. Accordingly (12.1) can be solved for s_r as a continuous function of s in the interval $a \leqq s \leqq a_1$. By Sturm's separation theorem it follows that P_r moves always in the same sense from Q_r to Q_{r+1} as P moves from Q to Q_1 in one sense. The lemma is thus proved. From Lemma 1 and the statement proved at the beginning of this section we have the following theorem:

THEOREM 8. *On a geodesic of Class IV the rth extremum of r following a given extremum of r is the rth conjugate point to that extremum. As a point P moves along the geodesic from one extremum to the next following, its rth conjugate point P_r moves continuously and in the same sense from the rth extremum following P to the $r+1$th.*

13. *Conjugate points on geodesics of Class V.*

THEOREM 9. *The angular measure, d, from a point on the outside equator to its first conjugate point is given by the formula*

$$d = \pi/(r_2 k)^{1/2}, \tag{13.1}$$

where k is the curvature of the generatrix [*cf.* (3.3)].

This theorem can be easily proved by setting up the Jacobi equation (8.4) for the outside equator and noting that $\bar{K}(s) = k/r_2$.

If the generatrix of the surface were an ellipse with axes $2a$ and [illegible] respectively parallel to the x and z coordinate axes one finds that

$$d = \pi b/(r_2 a)^{1/2},$$

and by taking a large enough relative to b, d can be made as small as [illegible]. Under the initial hypothesis [cf. (3.9)] $b < (2/3^{1/2})a$. Since $r_2 > 2a$ it follows that $d < b\pi/2^{1/2}a < (2/3)^{1/2}\pi$. *Thus d is never as great as* $(2/3)^{1/2}\pi$, *for the surface when the generatrix is an ellipse.*

PART IV. THE ENVELOPES OF THE GEODESICS.

14. *The integral of the equation of second variation.* Let there be given a geodesic of Class IV or V. Let the points neighboring this geodesic be referred to normal geodesic coordinates (s, n) (cf. § 8) so that the geodesic is the curve $n = 0$. Let s_0 be the s coordinate of a point P on the geodesic. Furthermore, let there be given an analytic family of geodesics neighboring $n = 0$ through P for which

$$n = \Phi(s, a), \qquad a \equiv \Phi_s(s_0, a). \tag{14.1}$$

Then

$$\Phi(s_0, a) \equiv 0, \qquad \Phi(s, 0) \equiv 0, \tag{14.2}$$

and the given geodesic $n = 0$ is given by $a = 0$. Let $\bar{w}(s)$ denote the solution of the Jacobi equation [cf. (8.4)] which satisfies the conditions

$$\bar{w}(s_0) = 0, \qquad \bar{w}'(s_0) = 1.$$

Denote by $z(s)$ the function $\Phi_{aa}(s, 0)$. Then $z(s)$ satisfies the differential equation

$$z'' + K(s, 0)z = -\bar{w}^2(\partial/\partial n)K(s, 0) \tag{14.3}$$

[cf. (8.5)], where the initial conditions $z(s_0) = z'(s_0) = 0$ obtain. These initial conditions follow from the identities in (14.1) and (14.2). In order to discuss the envelope of geodesics of Classes IV-V a formula for $z(s)$ at the rth conjugate point $(s_r, 0)$ to P on the geodesic $n = 0$ will be needed.

LEMMA 1. *At s_r, the rth conjugate point to s_0 on $a = 0$,*

$$z(s_r) = \Phi_{aa}(s_r, 0) = (-1)^r A \int_{s_0}^{s_r} \bar{w}^3(t) \frac{\partial K}{\partial n}(t, 0)\,dt \tag{14.4}$$

where A is a positive constant.

In order to prove this lemma one notes that

$$\bar{w}'' + K(s, 0)\,\bar{w} = 0. \tag{14.5}$$

Multiply (14.3) by $-\bar{w}$ and (14.5) by $+z$ and add the two equations. The result is the equation

$$z\bar{w}'' - \bar{w}z'' = \bar{w}^3\ (\partial/\partial n)K(s, 0).$$

Integrating from s_0 to s_r one obtains

$$z\bar{w}' - \bar{w}z' \Big]_{s_0}^{s_r} = \int_{s_0}^{s_r} \bar{w}^3(t) \frac{\partial K}{\partial n}(t, 0)\,dt.$$

At s_0 and its rth conjugate point s_r, $\bar{w}$ is zero. Moreover $z(s_0)$ is zero. Accordingly

$$z(s_r)\bar{w}'(s_r) = \int_{s_0}^{s_r} \bar{w}^3(t) \frac{\partial K}{\partial n}(t, 0)\,dt.$$

Now $\bar{w}'(s_r)$ is positive or negative according as r is even or odd. Hence formula (14.4) follows.

15. *The existence theorem for the rth envelope.* In order to establish an existence theorem for the rth envelope a lemma will be needed, concerning the zeros of a solution of the second order differential equation

$$u''(x) + p(x, \lambda)u'(x) + q(x, \lambda)u(x) = 0 \tag{15.1}$$

where p and q are analytic functions of x and λ in the domain

$$-d_1 \leqq \lambda \leqq d_1 \qquad d_1 > 0 \tag{15.2}$$

$$x_0 \leqq x \leqq x_0 + d_2 \qquad d_2 > 0. \tag{15.3}$$

It is known from the general theory of the second order differential equation that for any value of the parameter λ the zeros of a solution of this equation which is not identically zero are discrete. Consider those solutions $u(x,\lambda)$ such that for $x = x_0$ and any value of λ in (15.2)

$$\bar{u}(x_0,\lambda) = 0, \qquad \bar{u}_x(x_0,\lambda) = 1. \tag{15.4}$$

It will be assumed that for every value of λ in the interval (15.2) there are r values of x that make $\bar{u}(x,\lambda)$ zero in the interval (15.3). Order these r values of x for a given λ according to increasing magnitude and denote by $X(\lambda)$ the rth value of x thus obtained.

LEMMA 1. *The function $X(\lambda)$ which gives the rth zero of the solution $\bar{u}(x,\lambda)$ is a single valued analytic function of λ in the interval* (15.2).

This lemma can be proved from the well-known implicit function theorems. The proof is not given here. Take a point P with coordinates (u_0, v_0). Let there be given a system of geodesics on the surface expressible in the form

$$v = f(u,\beta) \tag{15.5}$$
$$\beta = f_u(u_0,\beta) \tag{15.6}$$

through P, where f is an analytic function of u and β for all values of u, and for β satisfying

$$-d \leqq \beta \leqq +d \qquad d > 0. \tag{15.7}$$

Let the rth conjugate points to P exist on all these geodesics for β in (15.7). Take

$$u_r = U_r(\beta), \qquad v_r = V_r(\beta) \tag{15.8}$$

as the u and v coordinates of the rth conjugate point to (u_0, v_0) on the geodesic β of (15.5).

THEOREM 10. *The functions $U_r(\beta)$ and $V_r(\beta)$ will be analytic for β in the interval* (15.7) *and will give an envelope to the system of geodesics* (15.5) [*or a point through which all the geodesics* (15.5) *pass*]. *Each of these geodesics will be tangent to the envelope at the rth conjugate point to P on that geodesic.*

Let the Jacobi equation be set up for each geodesic of the system (15.5), corresponding to the Calculus of Variations problem of minimizing the integral

$$[r^2(v) + (dv/du)^2]^{1/2}.$$

This equation is a second order differential equation expressible with coef

cients which are analytic functions of independent variable u and parameter β. Identify u and β of this Jacobi equation respectively with x and λ of (15.1). Lemma 1 then shows that $U_r(\beta)$ is an analytic function of β in the interval (15.7). Now substitute $U_r(\beta)$ for u in (15.5) and one finds that $V_r(\beta)$ is also analytic in the interval (15.7). That the curve thus determined by equations (15.8) satisfies the usual properties of the envelope as stated in the latter part of the theorem, or else gives a point through which pass all the geodesics of (15.5), follows as proved in Bolza for the "Im Kleinem" theorem (B. § 47(a), p. 358).

16. *The form of the envelope in the neighborhood of a conjugate point.* Let C be a given geodesic. For the sake of clearness in the geometric interpretation of the results take orthogonal geodesic coordinates (cf. § 8) such that $n = 0$ gives the geodesic C. The geodesic C will be called the central geodesic. The relative orientation of positive sense of measurement of s and n will be taken so that from the point of view of the external normal to the surface it will be the same as that of u and v, respectively, for a point on the outside equator. Take a point P on C with coordinates $(s_0, 0)$ and take a family of geodesics through P given by the equation

$$(16.1) \qquad n = \Phi(s, a)$$

where the parameter is chosen so that

$$(16.2) \qquad a = \Phi_s(s_0, a).$$

Then for the geodesic C, $a = 0$ (cf. § 14). Let P_r with coordinates $(s_r, 0)$ be the rth conjugate point to P on C (cf. § 8). Let P_r' with coordinates (s_r', n') be the rth conjugate point to P on any geodesic a' of system (16.1). The function Φ of (16.1) will be supposed analytic in the domain

$$(16.3) \qquad |a| \leqq d \qquad d > 0$$

$$(16.4) \qquad s_0 - e \leqq s \leqq s_r + e \qquad e > 0$$

where d is taken small enough so that $s = s_r'$ satisfies (16.4) when $a = a'$ satisfies (16.3).

In the last section it was found that the locus of the points P_r' could be written in the form

$$(16.5) \qquad s_r = f_1(a), \qquad n = f_2(a)$$

where f_1 and f_2 are analytic functions of a in the interval (16.3). In this section the possible forms that this envelope may take in the neighborhood of P_r will be discussed. Lindeberg* has investigated this problem for the

* *Mathematische Annalen*, Vol. 59 (1904), p. 321.

envelope in the neighborhood of the first conjugate point, and his results easily extended to apply to the rth envelope. Let

$$M = (\partial^m/\partial a^m)\Phi(s_r, 0) \tag{16.6}$$

the lowest order derivative of Φ with respect to a which does not vanish. It is found convenient to consider the possible forms which the envelope takes under five cases:

I. m is even and M negative.
II. m is even and M positive.
III. m is odd and M positive.
IV. m is odd and M negative.
V. M does not exist.

Case I. m is even and M negative. One denotes as the positive side of C near P_r the points neighboring C for which $n > 0$.

In the notation of this article equation (5) p. 324 of Lindeberg's article when applied to the rth envelope becomes*

$$n' = \Phi(s_r', a') - \Phi(s_r, 0) = (M/m - 2M)a'^m + R_1(a') \tag{16.7}$$

where, as defined above, s_r' is the s coordinate of a point on the rth envelope at which a curve $a = a'$ of family (16.1) touches it, and $R_1(a')$ denotes a power series in a' all of whose terms are of higher power than the mth. The series $R_1(a')$ is known to converge for values of a' in the interval (16.3). It follows from (16.7) that *in Case I the envelope lies entirely on the positive side of C in the neighborhood of P_r.*

For the rth envelope the relation which Lindeberg writes at the top of page 324 becomes under the notation of this article

$$s_r' - s_r = (-1)^{r-1}C_1 M a^{m-1} + R_2(a') \tag{16.8}$$

where C_1 is a positive constant and $R_2(a')$ is a power series in a' all of whose terms are of higher power than the $m - 1$th. $R_2(a')$ is known to converge in the interval (16.3). One replaces Lindeberg's $1/L^2$ by $(-1)^{r-1}C_1$ since it is clear from his definition of L^2 that its sign is positive or negative according as r is odd or even (cf. bottom of p. 322, Lindeberg, *loc. cit.*). From this formula it follows that *a neighboring geodesic to C of the family (16.1) through P determined by a' touches the envelope at a point P_r' for which*

* There is an error in (3) p. 323 of Lindeberg's article. It should read instead of N. This accounts for the replacing of $-N$ in Lindeberg's formula by $-2M$.

$s_r' - s_r$ is less than or greater than zero according as a' has the same sign as $(-1)^{r-1}$ or the opposite sign.

The formula of page 327 of Lindeberg's article becomes in the notation of this article

$$\bar{s} - s_r = (-1)^{r-1} C_2 M a'^{m-1}/m + R_3(a') \tag{16.9}$$

where C_2 is a constant greater than zero, $R_3(a')$ is of the same nature as $R_2(a')$ in (16.8), and where $\bar{s}$ gives the s coordinate of the intersection of the geodesic a' with the curve C in the neighborhood of P_r. This formula shows that *the neighboring geodesics to C through P for which a' has the same sign as $(-1)^{r-1}$ intersect C in the neighborhood of P_r at points between P and P_r while neighboring geodesics for which a' has the opposite sign intersect C on the far side of P_r from P.*

Case II. m is even and M positive. Here *the envelope near P_r is found to lie entirely on the negative side of C.* From (16.8) it is seen that $s' - s_r$ *has the same sign as $a'(-1)^{r-1}$ for neighboring geodesics to C through P.* Finally from (16.9) it is found that *neighboring geodesics to C through P for which $a'(-1)^{r-1} > 0$ intersect C in the neighborhood of P_r on the far side of P_r from P, and those for which $a'(-1)^{r-1} < 0$ intersect C between P and P_r near P_r.*

Case III. m is odd and M positive. Here it follows from formulae (16.7) and (16.8) that *the envelope has a cusp at P_r for which $s - s_r$ has the same sign as $(-1)^{r-1}$*, and that *a neighboring geodesic to C determined by a' in* (16.1) *touches the envelope at a point P_r' such that n' has the opposite sign to $a'(-1)^{r-1}$.* From (16.9) it follows that *neighboring geodesics to C intersect C near P_r at points on the farther or the nearer side of P^r from P according as $(-1)^{r-1}$ is positive or negative.*

Case IV. m is odd and M negative. Substitute $(-1)^r$ for $(-1)^{r-1}$ in the italicized statements under Case III and the results for Case IV are obtained.

Case V. m does not exist. Here *the equations* (16.5) *of the envelope give one point only,* and there is a system of geodesics through P and P_r.

17. *The rth envelope of geodesics of Classes IV-V through a point on the outside equator.* It has been seen that at a point on the outside equator the asymptotic geodesics cross this equator at an angle $\pm\chi_1$ (cf. § 7), whose cosine is given by the formula $\cos\chi_1 = r_1/r_2$. Geodesics of Class IV cross the outside equator at an angle χ, such that $\cos\chi = \alpha/r_2$, where α lies in the

interval $r_1 < \alpha < r_2$. Recall that Class V consists of just one geodesic; namely, the outside equator, and on this geodesic $\chi = 0$. If there be given a point P on the outside equator, and any value of χ in the interval

$$-\chi_1 < \chi < +\chi_1, \tag{17.1}$$

the geodesic G through P whose positive tangent makes such an angle χ with the positive tangent to the outside equator, will be a geodesic of Class IV or V. Furthermore all geodesics of Classes IV-V through P will be determined by these values of χ. Let the coordinates (u, v) of P be $(u_0, 0)$. All the geodesics through $(u_0, 0)$ except $u = u_0$ can be represented in the form

$$v = \psi(u, b), \qquad b = \psi_u(u_0, b), \tag{17.2}$$

where $\psi(u, b)$ is analytic in u and b for all values of u and b. In Differential Geometry * there is the relation

$$\tan \chi = (G/E)^{1/2} dv/du = (1/r)\, dv/du.$$

Thus in particular at $(u_0, 0)$, at which point $r = r_2$,

$$b = \psi_u(u_0, b) = r_2 \tan \chi. \tag{17.3}$$

Corresponding to the varying of χ on the interval (17.1), b by virtue of the preceding relation will vary on an interval

$$-b_1 < b < +b_1 \tag{17.4}$$

where

$$b_1 = r_2 \tan \chi_1.$$

Now for the investigation of the rth envelope of the system of geodesics given by (17.2) for b in the interval (17.4). It is known from Theorem 10 (cf. § 15) that this rth envelope is given by equations of the form

$$u_r = U_r(b), \qquad v_r = V_r(b) \tag{17.5}$$

where U_r and V_r are analytic functions of b for b in (17.4). In order to investigate this envelope consider any geodesic C of family (17.2)-(17.4) and take normal geodesic coordinates with C as the "central geodesic." In the neighborhood of this geodesic C, the geodesics can be given in the form

$$n = \Phi(s, a) \tag{17.6}$$

$$a = \Phi_s(s_0, a) \tag{17.7}$$

where $n = 0$, or $a = 0$ gives the geodesic C and $(s_0, 0)$ is the point P.

In the last section it was found that in investigating the envelope of

---

* Eisenhart, *loc. cit.*, p. 26 (24.)

the geodesics given by (17.6) that the lowest order derivative $(\partial^m/\partial a^m)\Phi(s,0)$ which does not vanish on that envelope plays a very important rôle. Preparatory to proving a theorem which applies the theory to the envelopes of geodesics of Class IV two lemmas will be proved. Take the "central geodesic" C of system (17.6)-(17.7) as a geodesic for which $b \neq 0$ when represented in the form (17.2). Let $\bar{P}_1$ denote the first intersection of C with the outside equator following P on C in the sense of increasing u. Let s be so measured on C that $s=0$ on C at $\bar{P}_1$ and so that the sense of increasing s is the sense of increasing u. Recall from the general theory (B. § 40) that for the analytic, positive, definite problem of the Calculus of Variations the relation between the surface coordinates (u, v) and the surface coordinates (s, n) is

$$(17.8) \qquad u = u(s,n), \qquad v = v(s,n), \qquad D(u,v)/D(s,n) \neq 0,$$

where u and v are analytic functions of s and n for points in the neighborhood of a finite segment of C.

LEMMA 1. *On C, $(\partial/\partial n)K(s,n)$ is less than or greater than zero according as the v coordinate of the point $(s,0)$ is greater or less than zero.*

Let $\bar{K}(u,v)$ be the surface curvature in terms of u and v. We can then write at the point given by (s,n) and (u,v)

$$\frac{\partial K}{\partial n} = \frac{\partial \bar{K}}{\partial v} \cdot \frac{\partial v}{\partial n} + \frac{\partial \bar{K}}{\partial u} \cdot \frac{\partial u}{\partial n}.$$

Denote by $\chi(s)$ the angle which the geodesic C makes with a parallel at a point $s=s$ on C [cf. (4.6)]. Then since $\partial\bar{K}/\partial u$ is zero at $(s,n)=(s,0)$,

$$(17.9) \qquad \frac{\partial K}{\partial n} = \frac{\partial \bar{K}}{\partial v} \cos\chi(s).$$

But it is known that $\partial u/\partial s$ is never zero on a geodesic of Class IV (cf. 4.5), so $\cos\chi(s)$ will always be positive. For $v>0$, $\partial\bar{K}/\partial v$ is negative and for $v<0$, $\partial\bar{K}/\partial v$ is positive. Thus the lemma follows.

Let the coordinates (s,n) of the rth conjugate point

$$P_r, \ (r = 1, 2, 3, \cdots, m)$$

to P on C be $(s_r, 0)$. Let M and M_1 be two successive points on C at which extrema of v occur and which include the point $\bar{P}_1$ between them. Then from Theorem 8 it is seen that P_1 *will lie between M and its conjugate M_1 on C.* The second lemma concerns a solution $\bar{w}(s)$ of the Jacobi equation,

$$(17.10) \qquad w''(s) + K(s,0)w(s) = 0,$$

which is zero at $s=s_0$ and which has first derivative equal to 1 there.

LEMMA 2. *If on the geodesic C emanating from P on $v = 0$ s be measured from $\bar{P}_1$, the first intersection of C with $v = 0$ following P, and if s_1 denote the s coordinate of the first conjugate point P_1 to P, then $\bar{w}(s)$, the solution of the Jacobi differential equation set up for C, vanishing at P, with its derivative equal to unity at P, satisfies the relation*

$$\bar{w}(-s) > \bar{w}(s) \tag{17.11}$$

for s in the interval

$$0 < s \leqq |s_1|. \tag{17.12}$$

In order to prove this lemma a function $w_1(s)$ is defined as follows

$$w_1(s) = \bar{w}(-s). \tag{17.13}$$

It is then noted that

$$w_1''(s) = \bar{w}''(-s) \tag{17.14}$$

and on account of the symmetry of the geodesic C and the surface with respect to the equatorial plane that

$$K(s, 0) = K(-s, 0). \tag{17.15}$$

Accordingly $w_1(s)$ will be a solution of the Jacobi equation (17.10), for if one sets $s = -s$ in (17.10) and employs (17.13), (17.14) and (17.15) one obtains

$$w_1''(s) + K(s, 0)w_1(s) = 0.$$

If one sets

$$W(s) = \bar{w}(s) - w_1(s) = \bar{w}(s) - \bar{w}(-s) \tag{17.16}$$

and proves that $W(s) < 0$ on the interval (17.12), Lemma 2 will be demonstrated. Now $W(s)$ is a solution of (17.10) and, by virtue of (17.13) $W(0) = 0$. From (17.16) and (17.13)

$$W(s_1) = -\bar{w}(-s_1) \tag{17.17}$$

$$W(-s_1) = \bar{w}(-s_1). \tag{17.18}$$

Using (17.17) if $s_1 > 0$ and using (17.18) if $s_1 < 0$ one obtains

$$W(|s_1|) < 0. \tag{17.19}$$

Since $W(s)$ vanishes at $s = 0$ and M and M_1 are conjugate points (cf. Theorem 8), $W(s)$ can vanish at no other point between M and M_1. But $W(s)$ is negative at $s = |s_1|$. $W(s)$ must accordingly be negative on the whole interval (17.12), and Lemma 2 is proved.

Consider the following theorem.

THEOREM 11. *On a geodesic of Class IV through a point P on the outside equator and for which $b > 0$, $\Phi_{aa}(s_r, 0)$ is greater or less than zero at the rth conjugate point $s = s_r$ according as r is odd or even.. If $b < 0$, the sign of $\Phi_{aa}(s_r, 0)$ is reversed in the above statement.*

From formula (14.4)

$$(17.20) \qquad \Phi_{aa}(s_r, 0) = (-1)^r A \int_{s_0}^{s_r} \bar{w}^3(t)\,(\partial K/\partial n)\,(t, 0)\,dt.$$

Now A is a positive constant. Take the central geodesic C as a geodesic of Class IV for which $b > 0$. Then the first part of Theorem 11 will be demonstrated if one shows that the integral

$$(17.21) \qquad \int_{s_0}^{s_r} \bar{w}^3(t)\,(\partial K/\partial n(t, 0)\,dt$$

is negative. Divide the proof into three principal headings; I when $r = 1$, II when $r = 2$, III when $r > 2$.

I. $r = 1$.

Consider the two cases: Case A, when P_1 lies on the outside equator or on the side of the equatorial plane for which $v > 0$; Case B, when P_1 lies on the side of the equatorial plane for which $v < 0$.

CASE A: *The first conjugate point P_1 is at $\bar{P}_1$ or between M and P_1.*

It appears from Lemma 1 that in this case $(\partial K/\partial n)\,(s, 0)$ is negative between P and P_1. But $\bar{w}(t)$ is positive in this interval. Thus when $r = 1$ for Case A the integral (17.21) is negative.

CASE B: *The first conjugate point P_1 is between P_1 and M_1 (s is measured from $\bar{P}_1$).*

On account of symmetry

$$K(-s, -n) = K(s, n).$$

Hence

$$(17.22) \qquad (\partial K/\partial n)\,(-s, 0) = -(\partial K/\partial n)\,(s, 0)$$

for all values of s. Write the integral (17.21) for $r = 1$ as

$$\int_{s_0}^{-s_1} \bar{w}^3(t)\,(\partial K/\partial n)\,(t, 0)\,dt + \int_{-s_1}^{0} \bar{w}^3(t)\,(\partial K/\partial n)\,(t, 0)\,dt + \int_{0}^{s_1} \bar{w}^3(t)\,(\partial K/\partial n)\,(t, 0)\,dt,$$

where s_1 is the value of s at P_1 and s_0 its value at P. The first of these integrals as in Case A is negative. In the third integral, from Lemma 2, $w(t) < w(-t)$ for t in the interval $0 < t \leqq s_1$. From formula (17.22)

...d Lemma 2 it is found that the third integral is positive and is l... ...solute value than the second integral, which is negative. Thus th... s... t... three integrals is negative and accordingly integral (17.21) is neg...tiv... ... the case under consideration.

... $r = 2$.

Let $\bar P_2$ denote the second intersection of the geodesic C with th... ...uator following P in the sense of increasing u, and let M_2 denote the ... ...int on C following M_1 at which v has an extremum. Write the int... ...7.21) for $r = 2$ in the form

$$(17.23) \qquad \int_{s_0}^{|s_1|} \bar w^3(t)\,\frac{\partial K}{\partial n}(t,0)\,dt + \int_{|s_1|}^{s_2} \bar w^3(t)\,\frac{\partial K}{\partial n}(t,0)\,dt.$$

... s_1 ... 0 from the discussion in I Case B it is known that the first int...gr... negative. If, however, $s_1 < 0$ one can apply Lemma 2 and the fact th... ...(t) is negative on the interval $0 < t < |s_1|$ to show that the first integr... ... negative just as was done in I Case B. In considering the second integral ...ove the origin of measurement of s to P_2 and as before let s_0 be the ... ...ordinate of P and s_r the rth conjugate point ($r = 1, 2, \cdots, m$). Con...id... ...o cases:

Case A, $s_2 \leqq 0$. Case B, $s_2 > 0$.

In terms of the arc length measured from $\bar P_2$ let $\bar s_1$ denote the value of s at ...he point whose s coordinate was $|s_1|$ when s was measured from $\bar P_1$. Th... ...e can write the second integral of (17.23) as

$$(17.24) \qquad \int_{\bar s_1}^{s_2} \bar w^3(t)\,\frac{\partial K}{\partial n}(t,0)\,dt.$$

CASE A: $s_2 \leqq 0$.

Here $\bar w(t)$ is negative for this integral except at s_2 and perhaps $\bar s_1$. Al... ...$K/\partial n(t,0)$ is positive for the integral except possibly at s_2 and $\bar s_1$. Th... ...e integral is negative.

...SE B: $s_2 > 0$.

Write the integral (17.24) as

$$\int_{\bar s_1}^{-s_2} \bar w^3(t)\,\frac{\partial K}{\partial n}(t,0)\,dt + \int_{-s_2}^{0} \bar w^3(t)\,\frac{\partial K}{\partial n}(t,0)\,dt + \int_{0}^{s_2} \bar w^3(t)\,\frac{\partial K}{\partial n}(t,0)$$

One treats these integrals as the three integrals of I Case B were treated. In the first integral $\bar{w}(t) < 0$ and $\partial K/\partial n > 0$ except possibly at the end values, and the integral is thus negative. A review of the proof of Lemma 2 will show that in II Case B $w(-s) < w(s)$ for the interval $0 < s \leqq |s_2|$. Accordingly the second integral above will be in absolute value larger than the third integral which is positive. The sum of all three will thus be negative.

III. $r > 2$.

Take $s = 0$ at the rth intersection $\bar{P}_r$ of C with the outside equator and consider the integral

$$\int_{\bar{s}_{r-1}}^{s_r} \bar{w}^3(t)\,(\partial K/\partial n)\,(t, 0)\,dt \tag{17.25}$$

where $\bar{s}_{r-1}$ is the s coordinate in terms of arc length measured from $\bar{P}_r$, of the point whose s coordinate in terms of arc length measured from $\bar{P}_{r-1}$ was $|s_{r-1}|$. If r is odd this integral can be proved negative by treating as in I; and if r is even it can be treated as the second integral of (17.23) was treated in II. If s_r is negative the integral

$$\int_{s_r}^{|s_r|} \bar{w}^3(t)\,(\partial K/\partial n)\,(t, 0)\,dt \tag{17.26}$$

also has to be considered, but this integral offers no new difficulty and can be shown to be negative by using Lemma 2 as in I if r is odd, and as in II with a reversal of sign, if r is even. Thus combining (17.25) and (17.26) it follows from I and II by mathematical induction that the integral (17.21) for $r > 2$ is negative when $b > 0$. If $b < 0$, from the symmetry of the situation as evidenced in formula (17.22), it is seen that the integral will be positive. Theorem 11 is thus demonstrated.

One can now prove the interesting theorem:

THEOREM 12. *The rth envelope to geodesics of Classes IV-V through a point P on the outside equator is given by equations* (17.5) *where* U_r *and* V_r *are analytic for* b *in the interval* (17.3). *It is not a point and is symmetric with respect to the equatorial plane of the surface. If r is odd; when* $b > 0$ *the envelope is tangent to the geodesics as in Case II,* § 16, *and when* $b < 0$ *it is tangent to the geodesics as in Case I. If r is even; when* $b > 0$ *Case I applies, and when* $b < 0$ *Case II applies.*

That such an envelope exists follows from Theorem 10. That it is no a point follows at once from Theorem 11 and § 16. Because of the symmetry of the surface and the geodesics issuing from P, the envelope will be symmetrical in the equatorial plane. The facts presented in the last two statements of the theorem are easily deduced from Theorem 11 and the "Du Kneinem" envelope theory of § 16.

23. *Further theorems concerning the rth envelope of geodesics through a point on the outside equator.* The following theorems can be proved by methods similar to those used in the foregoing sections. They are stated in the order in which their proofs would naturally be given.

THEOREM 13. *The rth envelope to geodesics of Classes IV-V through a point P on the outside equator has a cusp on the outside equator at the point $(u, 0)$ for which $u \geqq u_r$. The envelope has no other singularities. The vertex of the cusp divides the envelope into two branches on either of which v is an analytic function of u. On the branch of the envelope for which $v > 0$, $v' > 0$; and on the branch for which $v < 0$, $v' < 0$.*

THEOREM 14.

(a) *On a geodesic of Class IV the maximum at $r = \alpha$ of v increases continuously as b increases from 0 to b_1, where b_1 is the positive value of b corresponding to one of the asymptotic geodesics through P.*

(b) *The coordinate u of the first point of intersection with the outside equator of a geodesic of Class IV through a fixed point P on the outside equator increases continuously and without limit as b increases in the open interval from 0 to b_1.*

THEOREM 15. *The two branches of the rth envelope of geodesics of Class IV through a point $u = u_0$ on the outside equator for which $v > 0$ and $v < 0$ become asymptotic to the inside equator on the side for which $v > 0$ and $v < 0$ respectively.*

THEOREM 16. *For $v = \bar{v}$ let $\bar{u}_r$ and $\bar{u}_{r+1}$ be the corresponding u coordinates of points on the rth and $r + 1$th envelopes respectively. The difference $\bar{u}_{r+1} - \bar{u}_r$ increases continuously and without limit as v increases from 0 to the value of v on the inside equator.*

COROLLARY 1. *The rth conjugate point P_r to the point P on the outside equator for a geodesic of Class IV recedes from the outside equator as r is taken larger and larger.*

COROLLARY 2. *No two of the rth envelopes have a point in common.*

COROLLARY 3. *Let b_m denote the geodesic of Class IV which has its extreme values of v on the parallels $v = \pm \bar{v}$, and let $d(b_m)$ denote the difference of the u coordinates of successive extrema on b_m. As r becomes infinite, for v constant,* $\bar{u}_{r+1} - \bar{u}_r$ (see Th. 16) *decreases monotonically but always remains greater than $d(b_m)$.*

Somewhat similar theorems can be demonstrated concerning the envelopes of a system of geodesics of Class IV through a point not on the outside equator.

CORNELL UNIVERSITY.

Involutorial Transformations Belonging to a Linear Complex.

By Hannibal Albert Davis.

Introduction. Involutorial space transformations belonging to a linear complex* have been studied in considerable detail, especially by Montesano* and Pieri.† The methods employed by them are entirely synthetic. In the present paper the equations of the transformation belonging to a non-special linear complex are derived in the general case and in some of the particular cases. A number of new transformations are discussed. It has seemed advisable to include some of the results already known. The results of Sections 1, 4-6, 8, 10-12, 16-20 are due wholly or in part to Montesano, and those in Section 21 to Pieri.

I. *Transformations I Belonging to a Non-Special Linear Complex.*

1. *General Case.* Let the pairs of conjugate points P_1, P_2 in an involutorial space transformation I lie on the rays of a non-special linear complex Γ, a single pair on an arbitrary ray. The self-conjugate surface F in a congruence Q_1 of Γ, locus of the pairs of conjugate points situated on lines of the congruence, contains simply the directrices, p, q of Q_1, and passes through the pair of conjugate points on each line of Q_1, hence is of order four. A plane through p (or q) cuts F_4 in a residual cubic Δ_3 of genus 1, the self-conjugate curve in the Γ-pencil lying in the plane. Two cubics Δ_3 associated with two Γ-pencils having a line in common meet in the pair of conjugate points on the common line, hence the self-conjugate curve of a Γ-regulus R is a Δ_6 of genus 3 lying on the quadric containing R and having the lines of R as bisecants.

Since the base of a pencil of linear congruences is a regulus R, the pencil F_4 of associated self-conjugate surfaces has for base, in addition to the fundamental curve of the transformation, a Δ_6 of genus 3; hence the fundamental curve is a C_{10} of genus 11. Any plane π through an arbitrary

* "Su le trasformazioni involutorie dello spazio chi determinano un complesso lineare di rette," *Roma Accademia dei Lincei Rendiconti*, Ser. 4, Vol. 4 (1888), 207-215, 277-285.

† "Sulle trasformazioni birazionali dello spazio inerenti a un complesso lineare speciale," *Circolo Matematico di Palermo*, Vol. 6 (1892), pp. 240-244.

p not meeting C_{10} contains a Δ_3 associated with the Γ-pencil in π. The curves p and Δ_3 form the complete section of a surface F_4 by π, hence the fundamental curve C_{10} meets any plane in 10 points situated on a cubic curve Δ_3.

Since the image in I of a point P on C_{10} is a Δ_3 with a double point at P, the C_{10} is triple on the surfaces conjugate to the planes of space and the transformation is of order 11. The surfaces conjugate to the planes of space are $\Phi_{11}: C^3_{10}\ 20a_i$, where the a_i are parasite lines, quadrisecants of C_{10}.

The surface of invariant points is $K_8: C_{10}{}^2\ 20a_i$.

The jacobian of the transformation is the image surface $J_{40}: C_{10}{}^{11}20a_i{}^4$ of the fundamental C_{10}.

Since the conjugate in I_{11} of a point P on C_{10} is a $\Delta_3: P^2$, this Δ_3 meets C_{10} in 11 points. In general, the conjugate J of a fundamental curve C contains C to multiplicity equal to the number of points of intersection of C with the image curve of a point on it.

2. *The Equations of the Transformation.* There are $\infty^4 \mid F_4 \mid$, one for each linear congruence in Γ. Each F_4 of the web through a point P_1 passes also through a second point P_2 conjugate to P_1 in the I_{11}. But a single pencil $\mid F_4 \mid$ of these surfaces is sufficient to determine the transformation.

The directrices of a pencil $\mid Q_1 \mid$ of linear Γ-congruences form a regulus R', lying on the quadric H_2 with the regulus R, base of the pencil $\mid Q_1 \mid$. Through an arbitrary point P_1 of space passes one surface F_4 of the pencil $\mid F_4 \mid$ associated in I_{11} with the pencil $\mid Q_1 \mid$. The unique line from P_1 meeting the directrices of the congruence Q_1 associated with F_4 meets the F_4 in one other point P_2, conjugate to P_1.

Denote by

(1) $x_1x_3 - x_2x_4 = 0$, (2) $(x_1/x_2) = (x_4/x_3) = k$, (3) $(x_1/x_4) = (x_2/x_3) = m$,

the quadric H_2, the lines of R, and the lines of R' respectively. An F_4 which intersects (1) in a pair of lines (3) and a residual Δ_6 is of the form

$$m^2 f_1(k) + m f_2(k) + f_3(k) = 0,$$

where $f_i(k) = a_i k^4 + b_i k^3 + c_i k^2 + d_i k + e_i$, $(i = 1, 2, 3)$.
Two such surfaces are;

$$\begin{aligned} F \equiv\ & a_1x_1{}^4 + b_1x_1{}^3x_2 + c_1x_1{}^2x_2{}^2 + d_1x_1x_2{}^3 + e_1x_2{}^4 \\ & a_2x_1{}^3x_4 + b_2x_1{}^2x_2x_4 + c_2x_1x_2{}^2x_4 + d_2x_1{}^2x_2x_3 + e_2x_2{}^3x_3 \\ & + a_3x_1{}^2x_4{}^2 + b_3x_1x_2x_4{}^2 + c_3x_2{}^2x_4{}^2 + d_3x_2{}^2x_3x_4 + e_3x_2{}^2x_3{}^2 = 0, \\ F' \equiv\ & a_1x_1{}^2x_4{}^2 + b_1x_1{}^2x_3x_4 + c_1x_1{}^2x_3{}^2 + d_1x_1x_2x_3{}^2 + e_1x_2{}^2x_3{}^2 \\ & + a_2x_1x_4{}^3 + b_2x_1x_3x_4{}^2 + c_2x_1x_3{}^2x_4 + d_2x_1x_3{}^3 + e_2x_2x_3{}^3 \\ & + a_3x_4{}^4 + b_3x_3x_4{}^3 + c_3x_3{}^2x_4{}^2 + d_3x_3{}^3x_4 + e_3x_3{}^4 = 0. \end{aligned}$$

Each F_4 of the pencil

$$F - \rho F' = 0 \qquad (4)$$

contains two variable lines of R' and the Δ_6. Through a point $P_1(y)$ of space passes one F_4 of (4),

$$F - p^2 F' = 0, \qquad (5)$$

where $p^2 = F(y)/F'(y)$. Substituting F and F' in (5), we obtain

$$\begin{aligned}(6) \qquad & (a_1 x_1^2 + a_2 x_1 x_4 + a_3 x_4^2)\ (x_1^2 - p^2 x_4^2) + (b_1 x_1^2 + b_2 x_1 x_4 + b_3 x_4^2) \\ & (x_1 x_2 - p^2 x_3 x_4) + (c_1 x_1^2 + c_2 x_1 x_4 + c_3 x_4^2 + d_1 x_1 x_2 + d_2 x_1 x_3 \\ & + d_3 x_3 x_4 + e_1 x_2^2 + e_2 x_2 x_3 + e_3 x_3^2)(x_2^2 - p^2 x_3^2) = 0.\end{aligned}$$

This surface contains the two lines of R',

$$g_1 \equiv x_1 - p x_4 = 0,\ x_2 - p x_3 = 0;\ g_2 \equiv x_1 + p x_2 = 0,\ x_2 + p x_3 = 0.$$

The line r through $P_1(y)$ meeting g_1, g_2 meets them in $A(\alpha)$ and $B(\beta)$ respectively, where

$$\begin{aligned}&\alpha_1 = p(y_1 + p y_4),\ \alpha_2 = p(y_2 + p y_3),\ \alpha_3 = y_2 + p y_3,\ \alpha_4 = y_1 + p y_4, \\ &\beta_1 = -p(y_1 - p y_4),\ \beta_2 = -p(y_2 - p y_3),\ \beta_3 = y_2 - p y_3,\ \beta_4 = y_1 - p y_4.\end{aligned}$$

Any point on r has coordinates;

$$\begin{aligned}&\sigma x_1 = \lambda \alpha_1 + \mu \beta_1 = \tau(y_1 + p^2 n y_4), \\ &\sigma x_2 = \lambda \alpha_2 + \mu \beta_2 = \tau(y_2 + p^2 n y_3), \\ (7) \qquad &\sigma x_3 = \lambda \alpha_3 + \mu \beta_3 = \tau(y_3 + n y_2), \\ &\sigma x_4 = \lambda \alpha_4 + \mu \beta_4 = \tau(y_4 + n y_1),\end{aligned}$$

where $n = (\lambda + \mu)/p(\lambda - \mu)$. Substituting (7) in (6) and recalling that P, A, and B lie on (6), we get

$$\begin{aligned}(8) \qquad & [a_1 p^2 y_4 (2y_1 + p^2 n y_4) + a_2 (y_1^2 + p^2 n y_1 y_4 + p^2 y_4^2) \\ & + a_3 y_1 (n y_1 + 2 y_4)]\ [F' y_1^2 - F y_4^2] + [b_1 p^2 y_4 (2 y_1 + p^2 n y_4) \\ & + b_2 (y_1^2 + p^2 n y_1 y_4 + p^2 y_4^2) + b_3 y_1 (n y_1 + 2 y_4)]\ [F' y_1 y_2 - F y_3 y_4] \\ & + [c_1 p^2 y_4 (2 y_1 + p^2 n y_4) + c_2 (y_1^2 + p^2 n y_1 y_4 + p^2 y_4^2) \\ & + c_3 y_1 (n y_1 + 2 y_4) + e_1 p^2 y_3 (2 y_2 + p^2 n y_3) + e_2 (y_2^2 + p^2 n y_2 y_3 + p^2 y_3^2) \\ & + e_3 y_2 (n y_2 + 2 y_3) + d_1 p^2 (y_1 y_3 + y_2 y_4 + p^2 n y_3 y_4) \\ & + d_2 (y_1 y_2 + p^2 y_3 y_4 + p^2 n y_2 y_4) + d_3 (y_1 y_3 + y_2 y_4 \\ & \qquad + n y_1 y_2)]\ [F' y_2^2 - F y_3^2] = 0,\end{aligned}$$

where F and F' mean $F(y)$ and $F'(y)$. One notes that

$$\begin{aligned}&F' y_1^2 - F y_4^2 = [y_1 y_3 - y_2 y_4]\ [b y_1 y_4 + (c + d + e)(y_1 y_3 + y_2 y_4)], \\ &F' y_1 y_2 - F y_3 y_4 = [y_1 y_3 - y_2 y_4]\ [-a y_1 y_4 + (c + d + e) y_2 y_3], \\ &F' y_2^2 - F y_3^2 = [y_1 y_3 - y_2 y_4]\ [-a(y_1 y_3 + y_2 y_4) - b y_2 y_3],\end{aligned}$$

where $$a = a_1y_1^2 + a_2y_1y_4 + a_3y_4^2,\quad b = b_1y_1^2 + b_2y_1y_4 + b_3y_4^2,$$
$$c = c_1y_1^2 + c_2y_1y_4 + c_3y_4^2,\quad d = d_1y_1y_2 + d_2y_1y_3 + d_3y_3y_4,$$
$$e = e_1y_2^2 + e_2y_2y_3 + e_3y_3^2.$$

Equation (8) is now of the form

$$Pn + QF' = 0. \tag{9}$$

If we set

$$A = by_1y_4 + (c + d + e)(y_1y_3 + y_2y_4),$$
$$B = -ay_1y_4 + (c + d + e)y_2y_3,$$
$$C = -a(y_1y_3 + y_2y_4) - by_2y_3,$$

the coefficients in (9) become

$$\begin{aligned} P = {} & [a_1F^2y_4^2 + a_2FF'y_1y_4 + a_3F'^2y_1^2]\,A + [b_1F_2y_4^2 + b_2FF'y_1y_4 + b_3F'^2y_1^2]\,B \\ & + [c_1F^2y_4^2 + c_2FF'y_1y_4 + c_3F'^2y_1^2 + d_1F^2y_3y_4 + d_2FF'y_2y_4 + d_3F'^2y_1y_2 \\ & + e_1F^2y_3^2 + e_2FF'y_2y_3 + e_3F'^2y_2^2]C, \end{aligned}$$

$$\begin{aligned} Q = {} & [2(a_1F + a_3F')y_1y_4 + a_2(F'y_1^2 + Fy_4^2)]A + [2(b_1F + b_3F')y_1y_4 \\ & + b_2(F'y_1^2 + Fy_4^2)]\,B + [2(c_1F + c_3F')y_1y_4 + c_2(F'y_1^2 + Fy_4^2) \\ & + (d_1F + d_3F')(y_1y_3 + y_2y_4) + d_2(F'y_1y_2 + Fy_3y_4) \\ & + 2(e_1F + e_3F')y_2y_3 + e_2(F'y_2^2 + Fy_3^2)]C. \end{aligned}$$

Since $n = -F'Q/P$, the transformation is obtained from (7).

$$x_1 = Py_1 - FQy_4,\quad x_2 = Py_2 - FQy_3,\quad x_3 = Py_3 - F'Qy_2,\quad x_4 = Py_4 - F'Qy_1. \tag{10}$$

This is apparently of order 15, but there is a common factor $(y_1y_3 - y_2y_4)^2$. The factor $(y_1y_3 - y_2y_4)$ is contained once in both P and Q, and once more after the resulting expressions are combined in (10).

3. *Representations of I_{11} on a Double Space.* If the Plücker line coördinates are used, the complex to which I_{11} belongs may be written

$$\Gamma \equiv p_{12} - p_{34} = 0.$$

The lines of Γ may then be represented by the points of a quadric variety in 4-space by means of the transformation

$$p_{12} = p_{34} = \rho x_1,\ p_{13} = \rho x_2,\ p_{14} = \rho x_3,\ p_{42} = \rho x_4,\ p_{23} = \rho x_5, \tag{11}$$

$$x_1^2 + x_2x_4 + x_3x_5 = 0. \tag{12}$$

The projection of the points of this quadric from (0, 0, 1, 0, 0) into an ordi-

...y space $\sum_3'$ gives the Lie-Noether representation of the lines of Γ on ... points of $\sum_3'$,[*]

$$(13)\qquad p_{12} = p_{34} = \sigma x_1' x_3',\ p_{13} = \sigma x_2' x_3',\ p_{14} = -\sigma(x_1'^2 + x_2' x_4'),$$
$$p_{42} = \sigma x_3' x_4',\ p_{23} = \sigma x_3'^2,$$

$$(\ ')\qquad x_1' = \tau p_{12} = \tau p_{34},\ x_2' = \tau p_{13},\ x_3' = \tau p_{23},\ x_4' = \tau p_{42},$$
$$x_1^2 + x_2 x_4 + x_3 x_5 = 0.$$

This (1, 1) representation of the lines of Γ upon the points of $\sum$... ...nes a (1, 2) point transformation between the space $\sum_3$ of Γ and ... $\sum'$; the pair of conjugate points P_1, P_2 situated on a line of Γ co...respond ... the image point P' in $\sum_3'$ of the line.

The intersection with Γ of the complex

$$(14)\qquad a_1 p_{12} + a_2 p_{13} + a_3 p_{14} + a_4 p_{42} + a_5 p_{23} = 0$$

... linear congruence, whose image in $\sum_3'$ is

$$(a_1 x_1' + a_2 x_2' + a_5 x_3' + a_4 x_4') x_3' - a_3 (x_1'^2 + x_2' x_4') = 0.$$

Hence, $\infty^4 \mid F_4 \mid \sim \infty^4 \mid H_2' \mid :\ \gamma_2'$ where $\gamma_2' \equiv x_3' = 0,\ x_1'^2 + x_2' x_4' = 0$.

Associated with (0, 0, 1, 0, 0,) in $\sum_4$ is a line u of Γ which contains ... pair of conjugate points U_1, U_2.

$$u = p_{14} = 1,\ p_{ik} = 0,\ (ik \neq 14).$$

From (13), $x_3' = 0$, $x_1'^2 + x_2' x_4' \neq 0$. Therefore u (or U_1, U_2) $\sim$ a point of the plane $\alpha' - x_3' = 0$ not on the conic γ_2'.

Since the condition that the complex (14) contains u is $a_3 = 0$, ...s that $\infty^3 \mid F_4 \mid :\ U_1,\ U_2 \sim \infty^3 \mid H_2' \mid \equiv \infty^3 \mid \pi' \mid + \alpha'$, where the π' a... the planes of $\sum_3'$

It is easy to see that; $\infty^6 \mid \Delta_6 \mid \sim \infty^6 \mid C_2' \mid$, where $[C_2', \gamma_2'] = ...$ $\infty^4 \mid \Delta_6 \mid :\ U_1,\ U_2 \sim \infty^4 \mid r' \mid$, where the r' are the lines of $\sum_3'$ not meeting γ_2'. The remainder of the composite C_2' is a line in α' meeting ... $\infty^3 \mid \Delta_3 \mid \sim \infty^3 \mid \delta' \mid$, where the δ' are lines meeting γ_2'.

There are in a linear Γ-congruence Q_1, with distinct directrices p, q two series of Γ-pencils, one series for each directrix. The two series of Δ_3's belonging to the pencils correspond in $\sum'_3$ to the two sets of generators of the quadric ..., image of the F_4 associated with Q_1. If p coincides with q and ... belongs to Γ, the congruence Q_1 contains a single series of Γ-pencils. Th...

[*] See e. g. Sturm, *Liniengeometrie*, Vol. 1, pp. 257-269.

image in $\sum_3'$ of the F_4 is in this case a quadric cone through γ_2' whose vertix is the image of the pair of points P_1, P_2 on p, conjugate in the I_{11}.

Let P_1 in $\sum_3$ describe a line p, which together with its polar line q in Γ determines a linear Γ-congruence, hence an F_4. The image in $\sum_3'$ of p must be a curve C_n' on the H_2' which is the image of the F_4. A plane through p cuts F_4 in a residual Δ_3, self-conjugate curve in the I_{11} of the Γ-pencil in the plane. The Δ_3 meets p in three points and q in one point. Similarly, the Δ_3 in a plane through q meets p in one point and q in three points. Hence, the image in $\sum_3'$ of p (together with its conjugate c_{11} in I_{11}) is a C_4' of type [3, 1], while the image of q, polar of p in Γ, is a C_4' of type [1, 3] on the same quadric.

The image in $\sum_3'$ of a line p of Γ is a cubic C_3' on the cone K_2', image of the linear Γ-congruence with p for directrix. The C_3' passes through the vertex of K_2' and cuts each of its generators in one other point.

A Γ-congruence Q_n of degree n corresponds in $\sum_3'$ to a surface $F' : \gamma_2'^{\tau}$. Q_n contains n lines of each Γ-pencil and $2n$ lines of each Γ-regulus. The image of a Γ-pencil is a line δ' meeting γ_2', and the image of a Γ-regulus containing u is a line L' which does not meet γ_2'. Hence, $Q_n \sim F'_{2n} : \gamma_2'^n$, provided Q_n does not contain u.

In particular, as P_1 describes a plane π, the Γ-lines P_1P_2 form a congruence Q_n. An arbitrary plane α contains a Γ-pencil, hence a Δ_3, which cuts π in three points. The three lines of the Γ-pencil in α through these three points belong to Q_n, hence $n=3$. Therefore a plane (together with its conjugate Φ_{11} in I_{11}) corresponds in $\sum_3'$ to a surface $F_6' : \gamma_2'^3$, provided the plane contains neither U_1 nor U_2.

The surface of invariant points in I_{11} was seen to be $K_8 : C_{10}^2$.

$$[K_8, \Delta_6] = 48,\ [K_8, \Delta_3] = 24,\ [C_{10}, \Delta_6] = 20,\ [C_{10}, \Delta_3] = 10.$$

Since $\Delta_3 \sim$ a line meeting γ_2', and $\Delta_6 : U_1, U_2 \sim$ a line not meeting γ_2', it follows that the image $\sum_3'$ of K_8 is $L_8' : \gamma_2'^4$.

The image in $\sum_3'$ of the fundamental curve C_{10} is a ruled surface $R'_{20} : \gamma_2'^{10}$.

A few details of the Lie-Noether representation will now be considered. The corresponding details of the (2, 1) point transformation are easily supplied.

The linear Γ-congruence (Γ, u) composed of the lines of Γ which meet u is the intersection of Γ with the special linear complex $p_{23}=0$. From (13) it follows that $(\Gamma, u) \sim \alpha'^2$.

A point U on u has coördinates $(k, 0, 0, 1)$. A point in the polar plane

of U with respect to Γ is $(\alpha_1, \alpha_2, -k\alpha_2, \alpha_4)$. Hence, any line of the Γ-pencil with vertex U has coördinates: $p_{12} = k\alpha_2$, $p_{14} = k\alpha_4 - \alpha_1$, p_{42} [illegible] $p_{31} = -k^2\alpha_2$, $p_{23} = 0$, $p_{34} = k\alpha_2$. Equations (13) give as the image in Σ_3' the point $(k, -k^2, 0, 1)$ on γ_2' Hence, each line of a Γ-pencil meeting u corresponds to the same point on γ_2'.

The variable image of a linear Γ-congruence containing u was seen to be a plane π'. Such a congruence with distinct directrices p, q contains two series of pencils, which correspond to the two pencils of lines $[\pi' \gamma_2']$, while [illegible] represent the two pencils with vertices on u. But if $p = q$, the congruence contains a single series of pencils and a single pencil with vertex on u. The image plane π' is tangent to γ_2' at a point O', image of each line on the pencil on u.

An arbitrary Γ-pencil (A, α) contains a single line p meeting u in a point O. As A describes p, the image line δ' of (A, α) generates the pencil (O', π'), where π', image of the linear Γ-congruence with directrix p, is tangent to γ_2' at O', image of p. As p describes the Γ-pencil on O, the pencil (O', π') describes the bundle O'. As the point O describes u, the point O' describes γ_2', and we have the pencils of Γ represented by the lines meeting γ_2' as before noted.

4. If the fundamental C_{10} becomes composite in such a way as to have a ruled surface S_μ of parasitic lines, quadrisecants of the composite C_{10}, the surface S_μ is a part of each of the surfaces Φ, conjugates of planes in the I. Hence the transformation is of order $11 - \mu$. If the components of C_{10} are C_p, C_q, C_r, C_s of orders p, q, r, s, $p + q + r + s = 10$, and of multiplicities [illegible], 2, 3, respectively, on S_μ, the images of planes are $\Phi_{11-\mu}$: $C_p^3\, C_q^2\, C_r$ [illegible] a_k. The parasitic lines a are quadrisecants of the composite C_{10} which are not generators of S_μ.

The surface of invariant points is $K_{8-\mu}$: $C_p^2\, C_q\, a_1 \cdots a_k$. The C_s is a singular invariant curve of the transformation.

The lines of Γ which meet C_p form a congruence Q_p of degree p. The self conjugate surface F_{4p}: $C_p^{p+1}\, C_q^p\, C_r^p\, C_s^p$ in Q_p is composed of the image surface of J of C_p and the S_μ counted λ_1 times, where λ_1 is the number of points in which a generator of S_μ meets C_p. Hence the conjugate of C_p is

$$J_{4p-\lambda_1\mu}\colon C_p^{p+1}\, C_q^{p-\lambda_1}\, C_r^{p-2\lambda_1}\, C_s^{p-3\lambda_1}.$$

Similarly the conjugates of C_q, C_r and C_s are, respectively,

$$J_{4q-\lambda_2\mu}\colon C_p^q\, C_q^{q+1-\lambda_2}\, C_r^{q-2\lambda_2}\, C_s^{q-3\lambda_2},$$
$$J_{4r-\lambda_3\mu}\colon C_p^r\, C_q^{r-\lambda_3}\, C_r^{r+1-2\lambda_3}\, C_s^{r-3\lambda_3},$$
$$J_{4s-\lambda_4\mu}\colon C_p^s\, C_q^{s-\lambda_4}\, C_r^{s-2\lambda_4}\, C_s^{s+1-3\lambda_4},$$

Since $p+q+r+s=10$ and $\lambda_1+\lambda_2+\lambda_3+\lambda_4=4$, the sum of the orders of the surfaces J is $40-4\mu$; hence they form the complete jacobian of the transformation. The C_s being a singular invariant curve, $4s-\lambda_4\mu=0$. The multiplicity of each line a on J is equal to the number of points in which a meets that part of the fundamental curve associated with J.

The image in $\sum_3'$ of S_μ is in general a curve C_μ' which meets γ_2' in μ points. If the line u is a generator of S_μ, the image of S_μ is a $C_{\mu-1}'$ which meets γ_2' in $\mu-2$ points. In the former case $C_p \sim R_{2p}' : \gamma_2'^p$, $C_q \sim R_{2q}' : \gamma_2'^q$, $C_r \sim R_{2r}' : \gamma_2'^r$, $C_s \sim R'_{2s} : \gamma_2'^s$. The multiplicity of C_μ' on any R' is the number of points in which a generator of S_μ meets that part of the fundamental curve associated with the R'.

5. Suppose the surface S_μ is a pencil (A, α) of Γ. The self-conjugate surfaces in the net of linear Γ-congruences which contain the pencil (A, α) are, aside from the plane α, a net of cubic surfaces $|F_3|$. The self-conjugate curve of the Γ-pencil in a plane β through A is, aside from the parasitic line (α, β), a conic Δ_2 which, together with that part of the fundamental C_{10} not in α, forms the base of a pencil of the net $|F_3|$. Hence the C_{10} is composed of a C_7 of genus 5 through A and a γ_3 of genus 1 in α which meets the C_7 in the 6 points $[\alpha, C_7]$ other than A. The curve γ_3 is not the Δ_3 of α. The self-conjugate conic in any plane β through A meets C_7 in the 6 points $[\beta, C_7]$, other than A.

Through an arbitrary point P_1 of space passes a pencil of the net $|F_3|$, hence a Δ_2, self-conjugate in the Γ-pencil in a plane β through A. The line of Γ in β through P_1 meets Δ_2 in a second point P_2, conjugate to P_1 in the I_{10}.

The image in $\sum_3'$ of the net $|F_3| : C_7$ is the net $|H_2'| : \gamma_2'\delta'$, where δ' is the image of the pencil (A, α) in the Lie-Noether representation. The images of C_7 and γ_3 are, respectively, $R_{14}' : \gamma_2'^7\, \delta'$ and $R_6' : \gamma_2'^3\, \delta^3$. $[R_{14}', R_6'] = \gamma_2'^{21}\, \delta'^3\, 6\delta_i'\, C_{33}'$, where the $6\delta_i'$ are the images of the 6 Δ_2 associated in the I_{10} with the 6 points $[C_7, \gamma_3]$, and the C_{33}' is the image of the self-conjugate curve on the ruled surface R_{33} of lines meeting both C_7 and γ_3. The R_{33} does not include the pencil (A, α) nor the Γ-pencils through the 6 points $[C_7, \gamma_3]$.

6. Suppose the surface S_μ is a regulus R. The self-conjugate surfaces in the pencil of linear Γ-congruences whose base is R break up into the quadric S_2 of R and a pencil of quadric surfaces $|H|$. The fundamental C_{10} is composed of the C_4 of genus 1, base of the pencil $|H|$, and a hyperelliptic γ_6 of genus 3 on S_2 which meets C_4 in 8 points and which has for quadrisecants the lines of R.

Let the lines of the regulus R' of S_2, different from R, be $x_1/x_4 =$

$x_2/x_3 = m$. The complex Γ is cut by a pencil $A + tB = 0$ of linear complexes in a pencil of linear congruences whose self-conjugate surfaces are composed of S_2 and a pencil $H_1 + tH_2 = 0$ of quadric surfaces. With each quadric of the pencil associate a pair of lines of R' by means of the relation $m^2 = t$. Through a point P_1 of space passes one quadric H of the pencil. The line from P_1 cutting the pair of lines of R' associated with H meets H in a second point P_2 conjugate to P_1 in the I_9 thus defined. The equations of the I_9 are easily obtained and are free from extraneous factors.

The image in $\sum_3'$ of the pencil $|H|: C_4$ is the pencil of quadrics $|H_2'|:$ $\gamma_2' \ C_2'$, where C_2' is the image of the regulus R in the Lie-Noether representation. The images of C_4 and γ_6 are, respectively, $R_8' : \gamma_2'^4$ and R_{12}': $\gamma'^6 \ C'^2$. $[R_8'R_{12}'] = \gamma_2'^{24} \ 8\delta_i' \ C_{40}'$, where the $8\delta_i'$ are the images of the 8 Δ_2 associated in the I_9 with the 8 points $[C_4, \gamma_6]$; and the C_{40}' is the image of the self-conjugate curve on the ruled surface R_{40} of lines of Γ meeting both C_4 and γ_6 other than the 8 pencils through their points of intersection.

7. Suppose there are two pencils (A, α) and (B, β) of parasitic lines which do not have a line in common. The fundamental curves are the line $[\alpha, \beta] = k$, two conics C_2 in α and γ_2 in β, and a C_5 of genus 2. $[C_2, \gamma_2] = 0$, $[C_5, k] = 0$, $[C_5, C_2] = 4$, $[C_5, \gamma_2] = 4$.

The conjugates of planes in the I_9 are $\Phi_9 : C_5^3 \ C_2^2 \ \gamma_2^2 \ k \ 8a_i \ 2b_i$. The a_i are bisecants of C_5 meeting C_2 and γ_2, and the b_i are the trisecants of C_5 meeting k. These lines b_i are the only trisecants of C_5 which belong to Γ.

The surface of invariant points is $K_6: C_5^2 \ C_2 \ \gamma_2 \ 8a_i \ 2b_i$.

The conjugates in I_9 of the fundamental curves k, C_2, γ_2 and C_5 are, respectively, J_2: $C_5 \ 2b_i$, J_6: $C_5^2 \ C_2 \ \gamma_2^2 \ 8a_i$, J_6: $C_5^2 \ C_2^2 \ \gamma_2 \ 8a_i$, and J_{16}: $k^6 \ C_5^6 \ C_2^4 \ \gamma_2^4 \ 8a_i^2 \ 2b_i^3$. These surfaces form the jacobian of the transformation.

The pencils (A, α) and (B, β) are represented in $\sum_3'$ by the lines δ_1' and δ_2', respectively; and the images of the fundamental curves k, C_2, γ_2 and C_5 are, respectively, R_2': $\gamma_2' \ \delta_1' \ \delta_2'$, R_4': $\gamma_2'^2 \ \delta_1'^2$, R_4': $\gamma_2'^2 \ \delta_2'^2$, and R_{10}': $\gamma_2'^5$.

The surface J_2: C_5, together with the planes α and β, forms the F_4 associated in the I with the linear Γ-congruence which contains the pencils (A, α) and (B, β). J_2 is represented in $\sum_3'$ by $R_2' = \gamma_2' \ \delta_1' \ \delta_2'$, image of k.

Since a composite regulus is made up of two pencils with a common line, the transformation discussed here is quite distinct from that of Section 6.

8. Suppose the surface of parasitic lines is an S_3: k^2, where k is necessarily fundamental. A composite case is a regulus R and a pencil (A, α), the line k being part of the fundamental curve on each. The residual fundamental curves on R and α are, respectively, a γ_5 of genus 2 and a γ_2, each

meeting k in 2 points. Then γ_5 and γ_2 meet in 2 points. Therefore the fundamental curves are, besides the line k, a curve γ_7 of genus 3 on S_3 and a C_2 not on S_3. $[\gamma_7, C_2] = 6$. $[\gamma_7, k] = 4$. $[C_2, k] = 0$.

The images in $\sum_3'$ of the fundamental curves k, C_2, γ_7 are, respectively, R_2': $\gamma_2' \; C_3'$, R_4': $\gamma_2'^2$, R_{14}': $\gamma_2'^7 \; C_3'^3$, the C_3' being the image of S_3. Since the conjugate in I_8 of k is the singular plane β of C_2, the image of β is also R_2', which corresponds to the $K_4 \equiv S_3 + \beta$ of the general case.

$[R_2', R_4'] = \gamma_2' \; C_4'$, where the C_4' is the image of the ruled surface R_4 of Γ-lines meeting k and C_2.

$[R_2', R_{14}'] = \gamma_2'^7 \; C_3'^3 \; 4\delta_i' \; l'$ where the $4\delta_i'$ correspond to the image lines in I_8 of the 4 points $[k, \gamma_7]$, and l' corresponds to the curve Δ_3, image in I_8 of the single point $[C_7, l]$, the l being the simple directrix of S_3 and the polar in Γ of k.

$[R_4', R_{14}'] = \gamma_2'^{14} \; 6\delta_i' \; C_{22}'$, where the $6\delta_i'$ correspond to the image conics in I_8 of the 6 points $[C_2, \gamma_7]$, and C_{22}' corresponds to the self-conjugate curve on the R_{22} of Γ-lines meeting C_2 and γ_7.

9. Suppose there are a pencil (A, α) and a regulus R on H_2 of parasitic lines. If the intersection of α and H_2 is a proper conic γ_2, the resulting transformation is essentially different from that discussed in Section 8. The fundamental curves are the conic γ_2, a line k in α, a γ_4 of genus 0 on H_2, and a space C_3. The lines of R are trisecants of γ_4. $[\gamma_4, C_3] = 6$, $[k, \gamma_2] = 2$, $[k, \gamma_4] = 0$, $[k, C_3] = 2$, $[\gamma_4, \gamma_2] = 4$, $[\gamma_2, C_3] = 0$. The C_3 passes through A.

The conjugates of planes in the I_8 are $\Phi_8: C_3^3 \; \gamma_4^2 \; \gamma_2 k^2 \; 4a_i \; 3b_i$. The a_i are bisecants of C_3 which meet γ_2 and γ_4, and the b_i are bisecants of γ_4 which meet C_3 and k.

The surface of invariant points is $K_5: C_3^2 \; \gamma_4 \; k \; 4a_i \; 3b_i$.

The jacobian is composed of the conjugates in I_8 of the fundamental curves k, γ_2, C_3, and γ_4 which are, respectively, $J_3: \gamma_4 \; C_3 \; k \; 3b_i$, $J_4: \gamma_4 \; C_3^2 \; 4a_i$, $J_{11}: \gamma_4^3 \; C_3^4 \; \gamma_2^2 \; k^2 \; 4a_i^2 \; 3b_i$, and $J_{10}: \gamma_4^2 \; C_3^4 \; \gamma_2 \; k^4 \; 4a_i \; 3b_i^2$.

The images in $\sum_3'$ of (A, α) and R are, respectively, a line δ_1' meeting γ_2' and a conic C_2' meeting γ_2' in 2 points. The fundamental curves k, γ_2, C_3 and γ_4 (together with their conjugates in the I_8) correspond, respectively, to $R_2': \gamma_2' \; \delta_1'$, $R_4': \gamma_2'^2 \; \delta_1'^2 \; C_2'$, $R_6': \gamma_2'^3 \; \delta_1'$, and $R_8': \gamma_2'^4 \; C_2'^3$.

10. If the surface S_μ is an $S_4: k^2 l^2$, the double directrices are both necessarily fundamental. A composite form of such a surface is two quadrics $H_1: k \; l \; C_4$, $H_2: k \; l \; \gamma_4$, where the C_4 and γ_4 are of genus 1, and each meets both k and l in 2 points. Since $[C_4, \gamma_4] = 4$, the fundamental curve on S_4 is, in addition to k and l, a γ_8 of genus 5 which meets k and l each in 4 points.

The image in Σ_3' of S_4 is a curve C_4' of type [2, 2] on the quadric H_2' which corresponds in the Lie-Noether representation to the linear Γ-congruence with directrices k, l. To the fundamental lines k, l (of which each point on one corresponds in the I_7 to the whole of the other) correspond, respectively, the two reguli R_1', R_2' of H_2'. To γ_8 corresponds R_{16}': $\gamma_2'^8\ C_4'^2$.

$[H_2', R_{16}'] = \gamma_2'^8\ C_4'^2\ 8\delta_i'$, where the 8 δ_i' are the images in the Lie-Noether representation of the Γ-pencils with vertices at the 8 points of intersection of γ^8 with k and l.

11. If the surface S_μ is an S_4: C_3^2, a pair of quadrics H_1: $C_3\ \gamma_3$ [illegible] δ_3 may be taken as a composite case. $[C_3, \gamma_3] = 4$, $[C_3, \delta_3] =$ [illegible] $\delta_3] = 2$. It follows that the fundamental curve on S_4 is, in addition to C_3, a C_6 of genus 1 which meets C_3 in 8 points. The residual fundamental [illegible] on S_4 is a quadrisecant k of C_6.

The image in Σ_3' of S_4 is a C_4' of genus 0. The images of k, C_3 and C_6 are R_2': γ_2', R_6': $\gamma_2'^3\ C_4'^2$, and R_{12}': $\gamma_2'^6\ C_4'^2$, respectively.

12. Suppose in Section 7 the fundamental line k is on the quadric H_2, hence meets C_5 in 2 points, P and Q. Since P and Q are in α and β, the curves C_2 and γ_2 must pass through them. The lines of the Γ-regulus on H_2 are now parasitic, being trisecants of C_5 meeting k. The surfaces α, β, H_2 form a composite S_4: $k^3\ C_9$, where the $C_9 \equiv C_5 + C_2 + \gamma_2$ has triple points at P and Q and is of genus 4. A transformation I_7 results, in which k is a singular invariant line, the Δ_3 associated with the Γ-pencil on any point of it being composed of the 3 generators of S_4 through the point.

The image in Σ_3' of S_4 is a C_4' of type [3, 1] on the H_2': γ_2' which corresponds in the Lie-Noether representation to the linear Γ-congruence containing S_4. The images of k and C_9 are the regulus R' of trisecants to C_4' on H_2' and R_{18}: $\gamma_2'^9\ C_4'^3\ 2\delta_i'^3$.

13. Suppose we have a pencil (A, α) and a ruled surface S_3 of parasitic lines, and suppose α and S_3 intersect in a proper conic C_2 and a generator g_1 of S_3, which necessarily passes through A. The fundamental curves are the conic C_2, a line l in α, the double edge k of S_3, a curve C_5 of genus 1 on S_3 and a line m on neither S_3 nor α. $[C_5, C_2] = 3$, $[C_5, k] = 3$, $[C_5, l] = 1$, $[C_5, m] = 3$, $[C_2, k] = 1$, $[C_2, l] = 2$, $[C_2, m] = 0$, $[k, l] = 0$, $[k, m] = 0$, $[l, m] = 1$. The point $[C_5, l]$ is on g_1 and C_5 passes through A.

The conjugates of planes in the resulting I_7 are Φ_7: $C_5^2\ C_2\ k\ l^2\ m^3\ a$ [illegible] $2c_i$. The line a meets C_5, k, l and m; the b_i are bisecants of C_5 meeting C_2 and m; and the c_i are trisecants of C_5 meeting l.

The surface of invariant points is K_4: $C_5\ l\ m^2\ a\ 3b_i\ 2c_i$.

The jacobian is composed of the conjugates in I_7 of the fundamental curves C_5, C_2, k, l and m which are, respectively, $J_{13}: C_5^4\ C_2^2\ k\ l^4\ m^5\ a\ 3b_i^2\ 2c_i^3$, $J_3: C_5\ m^2\ 3b_i$, $J_1: l\ m\ a$, $J_3: C_5\ k\ l\ m\ a\ 2c_i$, and $J_4: C_5\ C_2\ k\ l\ m^2\ a\ 3b_i$.

14. Suppose there exist two reguli R_1 on H_1 and R_2 on H_2 of parasitic lines. We shall assume that no part of the $[H_1, H_2] = C_4$ is a line of Γ. The fundamental curves are the C_4, lines k_1, k_2 on H_1 and k_3, k_4 on H_2, and the two transversals l_1, l_2 of k_1, k_2, k_3, k_4. None of the k_i belong to either R_1 or R_2. $[C_4, k_i] = 2$, $[k_i, k_j] = 0$, $[C_4, l_i] = 0$, $[k_i, l_i] = 1$.

The conjugates of planes in the I_7 are $\Phi_7: C_4(k_1 \cdot\cdot k_4)^2(l_1 l_2)^3(a_1 \cdot\cdot a_4)$. The a_i are bisecants of C_4 meeting l_1 and l_2.

The surface of invariant points is $K_4: k_1 \cdot\cdot k_4\ l_1^2\ l_2^2\ a_1 \cdot\cdot a_4$.

The jacobian is composed of the conjugates in I_7 of the fundamental curves C_4, k_1, k_2, k_3, k_4, l_1 and l_2 which are, respectively $J_8: C_4(k_1 \cdot\cdot k_4)^2 (l_1 l_2)^4 (a_1 \cdot\cdot a_4)^2$, $J_2: k_1\ k_3\ k_4\ l_1\ l_2$, $J_2: k_2\ k_3\ k_4\ l_1\ l_2$, $J_2: k_1\ k_2\ k_3\ l_1\ l_2$, $J_2: k_1\ k_2\ k_4\ l_1\ l_2$, $J_4: C_4\ k_1 \cdot\cdot k_4\ l_1^2\ l_2\ a_1 \cdot\cdot a_4$, and $J_4: C_4\ k_1 \cdot\cdot k_4\ l_1\ l_2^2\ a_1 \cdot\cdot a_4$.

15. Given a C_5 of genus 1 and its surface of trisecants $R_5: C_5^2$. The R_5 belongs to a unique linear complex Γ'. There are ∞^4 cubic surfaces containing C_5. One surface F_3 of this system contains an arbitrary line l_1 of space. $F_3: C_5\ l_1\ l_2\ t_1 \cdot\cdot t_5$, where l_2 is the polar of l_1 in Γ' and $t_1 \cdot\cdot t_5$ are the trisecants of C_5 meeting l_1 and l_2. If l is in Γ' it is self-polar and $F_3: l\ t_1 \cdot\cdot t_5\ P^2$, where P is a point of l.* Conversely, for any point P of space there is a unique surface $F_3: C_5\ l\ t_1 \cdot\cdot t_5\ P^2$, where l is the only line of Γ' on F_3 through P. The other five lines of F_3 through P are bisecants of C_5. Associating the line l with the point P a (1, 1) correspondence is established between the lines of Γ' and the points of space, in which each line of Γ' is associated with a point on it. This correspondence will be denoted by M.

In M, a linear Γ-congruence with directrices l_1, $l_2 \sim F_3: C_5\ l_1\ l_2$. A Γ'-regulus $\sim C_4$, the residual base of a pencil of surfaces $F_3: C_5$. The C_4 is of genus 0 and meets C_5 in 10 points. A line of $R_5 \sim$ each of its points. Each line of the Γ'-pencil with vertex A on $C_5 \sim A$. An arbitrary Γ'-pencil $(A, \alpha) \sim$ a conic C_2 in α through A and through the five points $[\alpha, C_5]$.

In M^{-1}, a surface $F_k: C_5^h \sim$ a Γ'-congruence of order $(2k - 5h)$. A curve C_k having h points on $C_5 \sim$ a ruled surface $R_{3k-h}: C_k$. In particular, a

* See Colpitts, "On Twisted Quintic Curves," *American Journal of Mathematics*, Vol. 29 (1907), pp. 337-342; also Montesano, "Su la curva gobba di 5° ordine e di genere I," *Napoli Rendiconti*, Ser. 2, Vol. 2 (1888), pp. 181-188.

line l_1 not meeting $C_5 \sim R_3 : l_1\, l_2^2$, where l_2 is the polar of l_1 in Γ'. If l belongs to Γ', its image $R_3 : l^2$ is a Cayley scroll.

A surface F_3 of the system $\infty^4 \mid F_3 \mid : C_5$ determines a pair of lines l_1, l_2, polar in Γ'. A pencil of F_3 within the system determines a regulus of lines l_1, l_2, the associated regulus of which belongs to Γ'. A net of surfaces F_3 of the system determines a linear congruence of lines l_1, l_2, whose directrices are the Γ'-lines p_1, p_2 associated in M with the basis points P_1, P_2 of the net. Finally, a web of F_3 within the system determines a linear complex Γ of lines l_1, l_2. The complexes Γ' and Γ are in involution. If the web has a basis point, Γ is special, its directrix being the image in M^{-1} of the basis point.

Choose a web of $F_3 : C_5$ which does not have a basis point. A non-special linear complex Γ is then defined. A net of this web has two basis points P_1, P_2, either of which uniquely determines the other. For any point P_1 of space there exists such a net, hence a point P_2. Thus an I is established in space.

In M^{-1} a point P_1 of space corresponds to a line p_1 of Γ'. By the polarity in Γ, this line p_1 goes into a line p_2 which corresponds in M to P_2, conjugate to P_1 in the I. Since p_1, p_2 are polar in Γ, P_1P_2 is a line of Γ and I belongs to Γ.

A point P on C_5 corresponds in M^{-1} to the Γ'-pencil (P, π) which is transformed by the polarity in Γ into a Γ'-pencil (P', π'), where π' is the polar of P in Γ. Pencil (P', π') corresponds in M to a conic C_2 in π' through P. Hence C_5 is double on the surfaces Φ, conjugates of planes in I.

Each F_3 of the web is self-conjugate in I. Each curve C_4, base of a pencil in the web, is also self-conjugate in I. Since such a C_4 meets C_5 in 10 points, the transformation is of order 6.

The conjugate in I_6 of a generator t of R_5 is a point not in general on R_5. The locus of such points is a curve γ_5 of genus 1, simple on the surfaces Φ_6. $[C_5, \gamma_5] = 10$.

The conjugates of planes in I_6 are $\Phi_6 : C_5^2\, \gamma_5\, a_1 \cdot\cdot\, a_5$. The $a_1 \cdot\cdot\, a_5$ are trisecants of C_5 meeting γ_5. They are the only generators of R_5 which belong to Γ.

Let t_1 be a line of R_5. It, together with its conjugate T in I_6, belongs to a net of the web of F_3 defining Γ. A plane through t_1 meets C_5 in two points U and V, not on t_1. The line $TU = u$ has three fixed points on each F_3 of the net, hence there is within the net a pencil of F_3 which contains u. $\infty^1 \mid F_3 \mid : C_5\, u\, t_1\, t_2\, t_3$, where t_2 and t_3 are the trisecants of C_5 which meet u, other than t_1 and the two through U. The base of this pencil is a composite

[illegible]

Hence, the trisecants of γ_5 all meet C_5 and are parasitic in I_6. They form a surface S_5: $C_5\ \gamma_5^2$.

The point-wise invariant surface is K_3: $C_5\ a_1 \cdot \cdot\ a_5$, which belongs to the original system $\infty^4 \mid F_3 \mid : C_5$, but not to the web defining Γ.

The images in I_6 of the fundamental curves C_5 and γ_5 are, respectively, J_{15}: $C_5^5\ \gamma_5^3\ (a_1 \cdot \cdot\ a_5)^3$ and R_5: $C_5^2\ a_1 \cdot \cdot\ a_5$. These surfaces form the jacobian of the transformation.

16. If the surface S_μ is S_5: C_6^2, the fundamental curve on S_5 is, in addition to C_6, a space C_3. The remainder of the composite C_{10} is a line k, not on S_5. The self-conjugate surfaces in the ∞^2 quadratic Γ-congruences which contain S_5 are $\infty^2 \mid R_3 \mid : k^2\ C_3$.

Conversely, suppose we are given a net of surfaces $\mid R_3 \mid : k^2\ C_3$, where k is a bisecant of C_3, and a non-special linear complex Γ. Through a point P_1 of space passes a pencil of R_3 within the net. The base of this pencil, aside from the base of the net, is a pair of generators p_1, p_2 with p_1 through P_1. The plane polar to P_1 in Γ cuts p_2 in a point P_2. The I thus defined belongs to Γ. The surfaces of the net $\mid R_3 \mid : k_2\ C_3$ are self-conjugate in I, and each line of a pair p_1, p_2 is the conjugate of the other.

The simple directrices k' of the net $\mid R_3 \mid : k^2\ C_3$ form a congruence $Q_{2,\,3}$ of order 2 and class 3. Each line of Γ contains a single pair of points conjugate in I, except the lines of $Q_{2,\,3}$ which belong to Γ, each of which contains ∞^1 pairs of conjugate points. There exists a ruled surface S_5: $V^3\ C_3$ of such lines, where V, on k, is the vertex of the single cubic cone belonging to the net $\mid R_3 \mid : k^2\ C_3$. The surface S_5 is of genius 0, hence contains a double curve C_6: V^3. The transformation I is of order 6, and in it the conjugates of planes are Φ_6: $k^3\ C_3^2\ C_6\ a_1\ a_2\ a_3$ the a_1, a_2, a_3 being the bisecants of C_6 which meet k and C_3.

If we take for k the line $x_1 = 0$, $x_2 = 0$, and for C_3 the curve $x_1 = \lambda^2\mu$, $x_2 = \lambda\mu^2$, $x_3 = \lambda^3$, $x_4 = \mu^3$, the surfaces R_3: $k^2\ C_3$ are $(x_1^2 - x_2x_3)(a_1x_1 + a_2x_2) + (x_2^2 - x_1x_4)(a_3x_1 + a_4x_2) = 0$. Select a net of these by imposing the relation $\sum_1^4 \rho_i\, a_i = 0$. Through a point $P_1(y)$ of space passes a pencil of this net. Two arbitrary surfaces of the pencil have in common, besides the k and C_3, two lines p_1, p_2, of which p_1 passes through P_1. The polar plane of P_1 in $\Gamma \equiv p_{12} - p_{34} = 0$ cuts p_2 in a point P_2, whose coordinates give the equations of the transformation. These are readily obtained, and are of order 6, as they should be.

17. A ruled surface S_6: C_8^2 is contained in a net of quadratic congruences, each pencil of which has for residual base a regulus R which has

four lines in common with the S_6. There is a net of such reguli. An arbitrary ray of Γ lies on a single regulus of the net; but each generator of S_6 belongs to ∞^1 such reguli.

Each linear Γ-congruence of the pencil determined by a regulus R of the net has in common with S_6, besides the four lines $[R, S_6]$, a pair of generators. An involution J is thus set up on S_6, which is independent of the R. If h, k are a pair of conjugate lines in J, each linear congruence $Q_1: h, k$ contains a regulus R of the net, the generators of which establish a projectivity between the directrices of the congruence Q_1. These pairs of projective point ranges, which form a linear congruence with directrices h, k determine an I_5 belonging to Γ. The generators of S_6 are parasitic. The conjugates of planes in I_5 are $\Phi_5: h^2\, k^2\, C_8\, a_1 \cdot\cdot\, a_4$. The $a_1 \cdot\cdot\, a_4$ are bisecants of C_8 which meet h and k.

18. If a ruled surface $S_6: C_9{}^2$ belongs to a non-special linear complex Γ, there are two points, A and B, triple on S_6 and on C_9. The Γ-pencils (A, α) and (B, β), together with the surface S_6, form the base of a pencil of quadratic congruences of Γ. In any congruence Q_2 of this pencil is a pair of pencils whose planes π_1, π_2, intersect in the line k joining A and B. These planes π_1, π_2 are met by the lines of Q_2 in pairs of points. Thus a quadratic correspondence is established between the planes in which the triads of points in which the planes meet C_9, other than in A and B, are fundamental. As Q_2 describes the pencil of quadratic congruences, the planes π_1, π_2 describe the pencil of planes on k. The resulting ∞^1 quadratic correspondences determine in space an I which belongs to Γ. The generators of S_6 are parasitic lines, hence the transformation is of order 5. The conjugates of planes are $\Phi_5: k^3\, C_9\, a_1\, a_2$, where the a_1, a_2 are trisecants of C_9 which meet k.

19. Suppose that in the pencil of quadratic congruences of Section 18 there exists one, Q_2, formed by the lines of Γ which meet a conic C_2. The pair of planes π_1, π_2 which arises from Q_2 is composed of the plane ω of C_2 counted twice, the Γ-pencil (O, ω) being double in Q_2. Since the lines of (O, ω) are now parasitic, the transformation reduces to an I_4. The curve C_9 breaks up into the $C_2: A\, B$ and a $C_7: A^3\, B^2\, O$. The conjugates of planes are $\Phi_4: k^2\, C_7\, a$, the line a being a trisecant of C_7 which meets k.

20. If there exists a ruled surface S_4 of parasitic lines, the transformation I_4 can have no invariant surface. Since the congruence of Γ-lines determined by the invariant points [illegible] curve, triple on S_4, must be a C_4. The generators of S_4 are bisecants of C_4, which is therefore of genus 1.

[illegible]

of an arbitrary point P_1 with respect to the quadrics of the pencil meet in a line l. The polar plane of P_1 in Γ meets l in a point P_2. The pairs of points P_1, P_2 define an I which belongs to Γ. The C_4 is point-wise invariant in I. The congruence $Q_{2,6}$ of bisecants of C_4 contains a ruled surface $S_8: C_4^3$ of Γ-lines, each generator of which is parasitic. Hence the I is of order 3. The fundamental curve is, in addition to C_4, a C_6 of genus 3, double on S_8 and simple on the Φ_3, conjugates of planes in I_3. The jacobian of I_3 is $J_8: C_6^3$, conjugate to C_6.

No I of order < 3 can belong to a non-special linear complex.

II. *Transformations I Belonging to a Special Linear Complex.*

21. Suppose the pairs of conjugate points P_1, P_2 in an I lie on the rays of a special linear complex Γ, a single pair on an arbitrary ray. Let p denote the directrix of Γ.

The surface F_n, self-conjugate in a linear Γ-congruence with directrices p, q, contains simply the line q and meets an arbitrary line of the congruence in the pair of conjugate points lying on it. Hence if the surface is of order n, it contains the directrix p of Γ to multiplicity $(n-3)$. The n may be any integer greater than 3.

The self-conjugate surface in a bundle A of Γ is an $F_{n-1}: p^{n-4} A^{n-3}$. As the point A describes the line p, the surface F_{n-1} describes a pencil $| F_{n-1} |$, which determines the transformation. Through a point P_1 of space passes one surface $F_{n-1}: p^{n-4} A^{n-3}$ of the pencil. The line from P_1 through A meets F_{n-1} in one other point P_2. The pairs of points P_1, P_2 define the I belonging to Γ.

The self-conjugate curve in a Γ-pencil (A, α) with vertex A on p and plane α which does not contain p is a $\Delta_{n-1}: A^{n-3}$ of genus $(n-3)$. The self-conjugate curve in a Γ-pencil (B, β) with $\beta: p$ is a $\Delta_3: B$, of genus 1. The self-conjugate curve on a Γ-regulus is a C_{n+2} of genus $(n-1)$ which meets p in n points. Since a regulus is the base of a pencil of linear congruences, the fundamental curve, in addition to p, is a C_{5n-11} which meets p in $(5n-18)$ points. The curves p^{n-3} and C_{5n-11} form the base of a 4-fold system of surfaces F_n, hence the genus of C_{5n-11} is $(12n-38)$.

Each plane β through p is self-conjugate in I, the involution in it being a Geiser transformation having for fundamental points the 7 points $[\beta, C_{5n-11}]$ not on p. The I may be defined in this way.

A plane α through a line q skew to p and C_{5n-11} contains the Δ_{n-1} associated with the Γ-pencil in α. The curves q and Δ_{n-1} form the complete section of a surface F_n by α, hence any plane not through p meets C_{5n-11} in $(5n-11)$ points on a curve Δ_{n-1} which has an $(n-3)$-fold point on p.

The surface F_{n-1}: p^{n-4} A^{n-3} associated with the Γ-bundle with vertex A on p meets the first polar of A with respect to F_{n-1}, besides in p, in a curve C_{5n-11}, locus of the double points in I which are on F_{n-1}. The locus of coincidences in a Geiser involution is a sextic curve passing twice through each of the seven fundamental points. It follows that the point-wise invariant surface of the I is a K_{3n-4}: p^{3n-10} C^2_{5n-11}.

The conjugate in I of a plane α which does not contain p is a surface Φ which is cut by α in a point-wise invariant C_{3n-4} and a self-conjugate Δ_{n-1}. Hence the transformation is of order $(4n-5)$. The conjugates of planes in the I_{4n-5} are Φ_{4n-5}: p^{4n-13} $C_{5n-11}{}^3$ $(15n-40)a_i$. The parasitic lines a_i are trisecants of C_{5n-11} which meet p.

The conjugates in I_{4n-5} of the fundamental curves p and C_{5n-11} are, respectively, J_{4n-6}: p^{4n-14} $C_{5n-11}{}^3$ $(15n-40)a_i$, and J_{12n-18}: p^{12n-39} $C_{5n-11}{}^8$ $(15n-40)a_i{}^3$. These surfaces form the jacobian of the transformation.

22. As in the case of transformations I belonging to a non-special linear complex, the fundamental curve C_{5n-11} may become composite in such a way as to admit of ∞^1 trisecants meeting p which form a ruled surface S_μ. The I is then of order $(4n-5-\mu)$, the surface S_μ being a part of each of the surfaces Φ, conjugates of planes in the I. The surface of invariant points is also reduced in order to $(3n-4-\mu)$.

A simple directrix curve of S_μ may or may not be fundamental. A double directrix curve of S_μ is necessarily fundamental. No directrix curve other than p of multiplicity > 3 can exist on S_μ.

A few of the possibilities will be discussed.

23. Suppose the surface S_μ of parasitic lines is a Γ-pencil (A, α) with vertex A on p and plane α not containing p. The surfaces F_n associated in the transformation with the ∞^2 linear Γ-congruences which contain the pencil (A, α) break up into the plane α and a net of surfaces $|F_{n-1}|$, of which the base is the line p^{n-3} and that part of the fundamental C_{5n-11} not in the plane α. A pencil of the net associated with the linear Γ-congruences whose directrices form a pencil (M, α) has for residual base a conic Δ_2, which, together with the parasitic line MA, forms the self-conjugate Δ_3 in the plane (M, p). Hence the base of the net $|F_{n-1}|$ is, in addition to p, a C_{4n-10} of genus $(6n-20)$. The fundamental curve in α is then [illegible] of genus [illegible], which meets C_{4n-10} in $(4n-10)$ points. The C_{4n-10} meets p in $(4n-14)$ points.

The I_{4n-6} is completely determined by the net of surfaces ∞^2 $|F_{n-1}|$: [illegible]

The conjugates of planes in the I are Φ_{4n-6}: p^{4n-18} $C_{4n-10}{}^3$ $\gamma_{n-1}{}^2$ $(12n-33)a_i$. The a_i are bisecants of C_{4n-10} which meet γ_{n-1} and p.

The surface of invariant points is $K_{3n-5}: p^{3n-10}$ $C_{4n-10}{}^2$ γ_{n-1} $(12n-33)a_i$.

The jacobian of the transformation is composed of the conjugates in I of the fundamental curves p, C_{4n-10}, and γ_{n-1}, which are, respectively,

$$J_{4n-7}:\ p^{4n-14}\ C_{4n-10}{}^3\ \gamma_{n-1}{}^2\ (12n-33)\ a_i,$$
$$J_{8n-14}:\ p^{8n-26}\ C_{4n-10}{}^5\ \gamma_{n-1}{}^4\ (12n-33)\ a_i{}^2,\ \text{and}$$
$$J_{4n-7}:\ p^{4n-13}\ C_{4n-10}{}^3\ \gamma_{n-1}\ (12n-33)a_i.$$

24. The surface of parasitic lines may also be a pencil (B, α) with vertex B not on p and plane β through p. The base of the net of surfaces $|F_{n-1}|$ associated with the ∞^2 linear Γ-congruences which contain the pencil (B, β) is the line p^{n-4} and that part of the fundamental C_{5n-11} not in β. The residual base of a pencil of the net is a Δ_{n-2}. The fundamental curves are p, a space C_{5n-13} of genus $(12n-43)$ which meets p in $(5n-20)$ points, and a conic γ_2 in β which meets C_{5n-13} in six points. The seventh point $[C_{5n-13}, \beta]$ not on p is the vertex β of the pencil of parasitic lines.

The I_{4-n6} is determined by the net of surfaces F_{n-1} and the complex Γ.

The conjugates of planes in I are Φ_{4n-6}: p^{4n-14} $C_{5n-13}{}^3$ $\gamma_2{}^2$ $(15n-45)a_i$. The a_i are trisecants of C_{5n-13} which meet p.

The surface of invariant points is K_{3n-5}: p^{3n-11} $C_{5n-13}{}^2$ γ_2 $(15n-45)$ a_i.

The jacobian of the transformation is composed of the conjugates of the fundamental curves p and C_{5n-13}, which are, respectively,

$$J_{4n-7}:\ p^{4n-15}\ C_{5n-13}{}^3\ \gamma_2{}^2\ (15n-45)\ a_i\ \text{and}$$
$$J_{12n-21}:\ p^{12n-42}\ C_{5n-13}{}^8\ \gamma_2{}^6\ (15n-45)\ a_i{}^3.$$

The fundamental conic γ_2 is parasitic in the I, each point of it corresponding to the whole conic.

25. Suppose the composite C_{5n-11} admits of a regulus R of trisecants which meet p. Such a regulus, if composite, is made up of two pencils (A, α) and (B, β) with a common line l, one being one type, the other of the other, of those just discussed. Since the fundamental γ_{n-1} and γ_2 of the two pencils meet in the pair of conjugate points on l, the fundamental curve on the quadric S_2 containing R is, in addition to the line p, a γ_{n+1} of genus $(2n-6)$ which meets p in $(n-2)$ points. The residual fundamental curve, not on S_2, is a C_{4n-12} of genus $(6n-23)$ which meets p in $(4n-16)$ points and γ_{n+1} in $(4n-8)$ points.

The I_{4n-7} is determined by the pencil of surfaces $|F_{n-2}|$: p^{n-4} C_{4n-12} which is associated with the pencil of linear Γ-congruences containing R.

The conjugates of planes in I are Φ_{4n-7}: p^{4n-14} $C_{4n-12}{}^3$ $\gamma_{n+1}{}^2$ $(12n-36)a_i$. The a_i are bisecants of C_{4n-12} which meet γ_{n+1} and p.

The surface of invariant points is K_{3n-6}: p^{3n-11} $C_{4n-12}{}^2$ γ_{n+1} $(12n-36)$ a_i.

The jacobian is composed of the conjugates of the fundamental curves p, C_{4n-12} and γ_{n+1}, which are respectively,

$$\begin{aligned} J_{4n-8}&:\ p^{4n-15}\ C_{4n-12}{}^3\ \gamma_{n+1}{}^2\ (12n-36)\ a_i, \\ J_{8n-16}&:\ p^{8n-28}\ C_{4n-12}{}^5\ \gamma_{n+1}{}^4\ (12n-36)\ a_i{}^2, \\ J_{4n-8}&:\ p^{4n-14}\ C_{4n-12}{}^3\ \gamma_{n+1}\ \ (12n-36)\ a_i. \end{aligned}$$

26. Suppose the surface of parasitic lines is S_μ: $p^{\mu-1}$ q, $\mu \geqq 1$, where q is skew to p. It is easy to show by mathematical induction that the fundamental curves, aside from p are $\gamma_{n+2\mu-3}$ of genus $(2n+\mu-8)$ on S_μ which meets p in $(n+2\mu-6)$ points, and $C_{4n-2\mu-8}$ of genus $(6n-3\mu-17)$, not on S_μ, which meets p in $(4n-2\mu-12)$ points and $\gamma_{n+2\mu-3}$ in $(4n+2\mu-12)$ points.

A plane π through p meets $C_{4n-2\mu-8}$ in four points and $\gamma_{n+2\mu-3}$ in three points, aside from intersections on p. These seven points are fundamental points for a Geiser involution in π. As π describes the pencil p, an $I_{4n-\mu-5}$ is generated in which the conjugates of planes are $\Phi_{4n-\mu-5}$: $p^{4n-\mu-12}$ $C_{4n-2\mu-8}{}^3$ $\gamma_{n+2\mu-3}{}^2$ $(12n-3\mu-30)a_i$. The a_i are bisecants of $C_{4n-2\mu-8}$ which meet $\gamma_{n+2\mu-3}$ and p.

The surface of invariant points is $K_{3n-\mu-4}$: $p^{3n-\mu-9}$ $C_{4n-2\mu-8}{}^2$ $\gamma_{n+2\mu-3}$ $(12n-3\mu-30)$ a_i.

The jacobian of the transformation is composed of the conjugates of the fundamental curves p, $C_{4n-2\mu-8}$ and $\gamma_{n+2\mu-3}$ which are, respectively:

$$\begin{aligned} J_{4n-\mu-6}&:\ p^{4n-\mu-13}\ C_{4n-2\mu-8}{}^3\ \gamma_{n+2\mu-3}{}^2\ (12n-3\mu-30)\ a_i, \\ J_{8n-2\mu-12}&:\ p^{8n-2\mu-24}\ C_{4n-2\mu-8}{}^5\ \gamma_{n+2\mu-3}{}^4\ (12n-3\mu-30)\ a_i{}^2, \\ J_{4n-\mu-6}&:\ p^{4n-\mu-12}\ C_{4n-2\mu-8}{}^3\ \gamma_{n+2\mu-3}\ (12n-3\mu-30)\ a_i. \end{aligned}$$

27. If two Γ-pencils (B, β) and (B', β') with vertices not on p are parasitic, the ∞^3 conics which contain B and B' and meet p are the directrices of ∞^3 Γ-congruences $Q_{1,2}$ containing (B, β) and (B', β'). The associated web of self-conjugate surfaces is, aside from β and β', a web $|F_{n-1}|$: p^{n-4} C_{5n-15}. If, in particular, the fundamental conics γ_2 in β and γ_2' in β' intersect in a pair of points O_1 and O_2, the C_{5n-15} meets p in $(5n-20)$ points, is of genus $(12n-46)$, passes through B and B', and meets γ_2 and γ_2' each in four points. The points O_1 and O_2 are $(n-3)$-fold basis points in the web F_{n-1}. Now set $n=4$ and we have a well known I_9.

See Sharpe and Snyder, "Certain Types of Involutorial Space Transformations," *Transactions of the American Mathematical Society*, Vol. 21 (1920), p. 58. Also [illegible] (1923), pp. [illegible].

On the Inverse Problem in the Calculus of Variations.

Thomas H. Rawles.

In a previous paper we have applied the theory connected with the invariant integral to the inverse problem in the calculus of variations.* The plan of attack was to determine in the first place the general form of integrand function which results from a prescribed relation between the slope of the extremal and the normal to the transversal curves or surfaces. This determination involves an arbitrary function of the coördinates which in turn may be determined by assigning a family of curves as the extremals of the problem. Conditions were given under which an arbitrary family of curves may be taken as the extremals of a problem involving a given transversality.

The object of the present paper is to determine directly the most general form of the integrand function which may have as extremals a given two parameter family of curves, a problem which originates with Darboux.† By approaching this problem from the standpoint of the invariant integral we can obtain the solution by somewhat simpler processes than those involved in the method of Darboux.

We have first to show that if $y(x, a, b)$ is a general solution of the Euler equation arising from a problem in two dimensions, a and b being any two constants whatever, we can construct a function having as arguments x, y, and one of these constants, $W(x, y, a)$ say, such that $(\partial/\partial a) W(x, y, a) = b'$ is a general solution of the Euler equation.

Such a function may be obtained from the definite integral

$$(1) \qquad J(x, a, b) = \int_X^a f\,[x, y(x, a, b), y_x(x, a, b)\,]\,dx,$$

where $f(x, y, y')$ is the integrand of the integral minimized by the curves $y = y(x, a, b)$. The lower limit, X, is itself a function of a and b such that the equations $x = X(a, b)$, $y = Y\,[X(a, b), a, b]$ represent a transversal curve associated with $y = y(x, a, b)$, a being assigned a fixed value, and b being taken as the parameter.

Now if $y_b(x, a, b) \neq 0$ for a certain set of values of the arguments we

* Rawles, *Transactions of the American Mathematical Society*, Vol. 30, pp. 765-784.

† Darboux, *Théorie des Surfaces*, Vol. III, paragraphs 604, 605, 606.

may solve the equation $y = y(x, a, b)$ for b and obtain $b = B(x, y, a)$. When this result is substituted in the left member of (1) the result is

$$(2) \qquad J[x, a, B(x, y, a)] = W(x, y, a).$$

We calculate the derivatives of the function W by differentiating (1) and obtain the familiar expressions

$$(3) \qquad \partial W/\partial x = f(x, y, p) - p f_{y'}(x, y, p), \; \partial W/\partial y = f_{y'}(x, y, p).$$

where $p(x, y, a) = y_x[x, a, B(x, y, a)]$.

In the same manner the derivatives of W_a are given by

$$(4) \qquad \partial W_a/\partial x = - p f_{y'y'}(x, y, p) p_a, \; \partial W_a/\partial y = f_{y'y'}(x, y, p) p_a.$$

We now consider the curve defined by

$$(5) \qquad W_a = b'.$$

Along this curve we must have $f_{y'y'} p_a(dy - dx\, p) = 0$. If $f_{y'y'} \equiv 0$ the integrand function contains the derivative y' only as a linear term. This would indicate a degenerate form of the problem which we do not consider. Also $p_a \not\equiv 0$; for if it were $y(x, a, b)$ would not be a general solution of a second order differential equation.*

Limiting ourselves, then, to those values of x, y, and y' for which $f_{y'y'} \neq 0$, $p_a \neq 0$, it follows that along the curve defined by (5) $dy/dx = p(x, y, a)$.

By the differentiation of (2) we can actually calculate W_a. We find

$$W_a = - f_{y'} Y_a - [f - y_x(X, a, b) f_{y'}] X_a,$$

where the arguments of f and $f_{y'}$ are X, Y, and $y_x(X, a, b)$ and b is replaced by $B(x, y, a)$. When the function W_a reduces simply to $B(x, y, a)$ the constants a and b are said to be canonical. In general, however, we obtain $W = \phi\,[a, B(x, y, a)]$; and it can be shown further that, under the assumptions made, $\partial\phi/\partial b \neq 0$.

The theory which we have outlined leads us to a method of obtaining the integrand functions which are minimized by a given two parameter family of curves. Let us take as our extremals the family represented by $y = y(x, a, b)$ and solve for b finding $b = B(x, y, a)$. We now form an arbitrary function, $\phi\,[a, B(x, y, a)]$, which may be regarded as the derivative of the transversal function, $W(x, y, a)$.

When we integrate with respect to a we obtain

$$(6) \qquad W(x, y, a) = \int \phi[a, B(x, y, a)]\, da + \theta(x, y),$$

* Bolza, [illegible] p. 148.

where θ is an arbitrary function. Also we may differentiate $B(x, y, a) = b$ with respect to x considering y as a function of x, and then solve for a, finding $a = A(x, y, p)$.

On the other hand we see by (3) that

$$f(x, y, p) = \partial W/\partial x + p\, \partial W/\partial y, \tag{7}$$

where a is replaced by $A(x, y, p)$ after the differentiation. This gives us finally,

$$\begin{aligned} f(x, y, p) = (\partial/\partial x) \int \phi[a, B(x, y, a)]da + p(\partial/\partial y) \int \phi[a, B(x, y, a)]da \\ + \theta_x(x, y) + p\theta_y(x, y), \\ a = A(x, y, p). \end{aligned} \tag{8}$$

Conversely, to show that (8) has as its extremals the given family of curves we have only to reverse the argument. If we form the Hamiltonian equation associated with (8) we have in (6) a solution already at hand.* It then follows that the equations of the extremals are given by

$$(\partial/\partial a) W(x, y, a) = \phi\, [a, B(x, y, a)] = b'.\dagger$$

Finally, to restore the original constants we put $b' = \phi(a, b)$ and the equations of the extremals assume the form

$$B(x, y, a) = b. \tag{9}$$

To illustrate the method we shall take the problem of determining the integrand function which has as its extremals straight lines. If the family of extremals is $y = ax + b$, $B(x, y, a) = y - ax$, and $a = p$. For the integrand function we find

$$\begin{aligned} f(x, y, p) = (\partial/\partial x) \int \phi(a, y - ax)\, da + p(\partial/\partial y) \int \phi(a, y - ax)\, da \\ + \theta_x(x, y) + p\theta_y(x, y), \\ a = p. \end{aligned} \tag{8a}$$

If we differentiate with respect to x and y under the integral signs (8a) becomes

$$\begin{aligned} f(x, y, p) = \int (p - a)\phi'(a, y - ax)\, da + \theta_x(x, y) + p\theta_y(x, y), \\ a = p, \end{aligned}$$

where ϕ' indicates the derivative of ϕ with respect to the second argument. This is the form of solution given by Darboux.‡

---

* Bolza, *loc. cit.*, p. 132.

† Bolza, *loc. cit.*, p. 138.

‡ Darboux, *loc. cit.*, paragraph 606; also, Bolza, *loc. cit.*, p. 39.

An Application of the Laguerre Method for the Representation of Imaginary Points.

By B. M. Turner.

1. *Introduction.* In an earlier paper * the writer directed attention to the fact that while three collinear real points of inflexion impose but five conditions on a real † non-singular plane cubic curve, and hence leave the curve with four degrees of freedom; still not one of the six imaginary points of inflexion may be chosen arbitrarily. The statement of the fact was followed by a discussion of the positions of the imaginary points of inflexion and critic centers for the four-fold infinite system of cubics. This paper shows that the variable imaginary inflexions and critic centers, represented by real point-pairs in accordance with the Laguerre method for the representation of imaginary points, describe unique systems of curves; and brings out more clearly the relations of the sets of points.

2. *The Laguerre Method.*‡ In the Laguerre representation of points in a plane the line at infinity, $z = 0$, and the circular points

$$(1, i, 0), \quad (1, -i, 0), \qquad\qquad i = (-1)^{1/2},$$

are fixed. Two lines, extending into the finite part of the plane and passing through the circular points in the order given, are called the first and second minimal lines of their intersection. Each real point is represented by itself; and each finite imaginary point by the pair of real points which lie on its minimal lines, the first and second point of the pair lying respectively on the first and second minimal line. The same pair of points in the reverse order represents the conjugate imaginary point.

For a pair of imaginary points the intersections of the first minimal line of each one by the second minimal line of the other form the representative real point-pair. For the imaginary point $(a + bi,\ c + di,\ 1)$ the minimal lines are

$$\begin{vmatrix} x & y & z \\ a + bi & c + di & 1 \\ 1 & i & 0 \end{vmatrix} = 0:$$

* *Annals of Mathematics*, Vol. 33, No. 4 (June, 1922).

† [illegible]

‡ [illegible]

and the real points on these lines $(a-d,\ c+b,\ 1)$, $(a+d,\ c-b,\ 1)$ form the representative point-pair. The pair in the opposite order

$$(a+d,\ c-b,\ 1),\ (a-d,\ c+b,\ 1)$$

represent the conjugate point $(a-bi,\ c-di,\ 1)$.

By definition a chain * is a system of collinear points which satisfy the following conditions:

(1) The cross-ratios of any four are real.

(2) With three arbitrarily chosen points of the system there exists a fourth which makes any real cross-ratio other than zero, one and infinity.

In the Laguerre representation a chain appears as two real lines or circles.

3. *Chains of Inflexions.* By the earlier paper the system of real cubics with $I_1\ (0,1,-1)$, $I_2\ (-1,0,1)$, $I_3\ (1,-1,0)$ as points of inflexion and $\Sigma\ (1,1,1)$ as fixed critic center common to the three real harmonic polars, has an equation

$$\begin{aligned}(x+y+z)(x^2+y^2+z^2+yz+zx+xy)&\\ +\lambda(\alpha x+y+z)(x+\alpha y+z)(x+y+\alpha z)&=0.\end{aligned}$$

The real harmonic polars are

$$h_1\equiv y-z=0,\quad h_2\equiv z-x=0,\quad h_3\equiv x-y=0;$$

and the imaginary inflexions are

$$\begin{array}{ll}(\omega^2-\omega,\ \ 1-\alpha\omega^2,\ \ \alpha\omega-1), & (\omega-\omega^2,\ \ 1-\alpha\omega,\ \ \alpha\omega^2-1);\\ (\alpha\omega-1,\ \ \omega^2-\omega,\ \ 1-\alpha\omega^2), & (\alpha\omega^2-1,\ \ \omega-\omega^2,\ \ 1-\alpha\omega);\\ (1-\alpha\omega^2,\ \ \alpha\omega-1,\ \ \omega^2-\omega), & (1-\alpha\omega,\ \ \alpha\omega^2-1,\ \ \omega-\omega^2);\end{array}$$

where $1, \omega, \omega^2$ are the cube roots of unity.

The cross-ratios of the four points

$$(\omega^2-\omega,\ \ 1-\alpha_j\omega^2,\ \ \alpha_j\omega-1),\qquad (j=1,2,3,4),$$

are those of the group containing

$$(\alpha_1-\alpha_2)(\alpha_3-\alpha_4)/(\alpha_1-\alpha_4)(\alpha_3-\alpha_2),$$

all real. With three points

$$(\omega^2-\omega,\ \ 1-\alpha_j\omega^2,\ \ \alpha_j\omega-1),\qquad (j=1,2,3),$$

* Von Staudt, *Beiträge zur Geometrie der Lage*, Part II, Nüremburg (1858), pp. 137 ff. Coolidge, *loc. cit.*, p. 36.

there exists a fourth which makes a real cross-ratio r; namely,

$$(\omega^2 - \omega, \quad 1 - \alpha_4\omega^2, \quad \alpha_4\omega - 1),$$

with $\qquad \alpha_4 = [r\alpha_1(\alpha_3 - \alpha_2) + \alpha_3(\alpha_2 - \alpha_1)] / [r(\alpha_3 - \alpha_2) + (\alpha_2 - \alpha_1)].$

Then as α varies, the points $(\omega^2 - \omega, 1 - \alpha\omega^2, \alpha\omega - 1)$ form a chain on the line $x + \omega y + \omega^2 z = 0$, and similarly five other chains are described. The totality is three pairs of chains, the mates of each pair lying one on each of the lines

$$x + \omega y + \omega^2 z = 0, \text{ and } x + \omega^2 y + \omega z = 0,$$

which are equiharmonic to the three real harmonic polars. Hence *the imaginary points of inflexion for the doubly infinite system of real cubics, determined by three real points of inflexion and the critic center common to the three real harmonic polars, form three pairs of chains, the mates of each pair lying one on each of the Hessian lines of the real harmonic polars.*

With the fixed line at infinity, $z = 0$, I_1 and I_2 are finite points and I_3 the point at infinity on the line

$$I_1I_2 \equiv x + y + z = 0.$$

By the Laguerre method for the representation of imaginary points in a plane, each pair of the imaginary points of inflexion is represented by a pair of real points. The pair

$$(\omega^2 - \omega, 1 - \alpha\omega^2, \alpha\omega - 1), \quad (\omega - \omega^2, 1 - \alpha\omega, \alpha\omega^2 - 1)$$

is represented by

$$\{\pm 3^{1/2}\alpha^2 + [3 \pm 2(3^{1/2})]\alpha, \quad -\alpha^2 + (2 \pm 3^{1/2})\alpha + 2 \pm 2(3^{1/2}), \quad -2(\alpha^2 + \alpha + 1)\},$$

the upper signs throughout giving one of the real points and the lower signs the other.* As α varies, these real points describe the two circles

$$x^2 + y^2 - xz \pm 3^{1/2}yz - (1 \pm 3^{1/2})z^2 = 0.$$

Similarly the pair of imaginary inflexions

$$(\alpha\omega - 1, \omega^2 - \omega, 1 - \alpha\omega^2), \quad (\alpha\omega^2 - 1, \omega - \omega^2, 1 - \alpha\omega)$$

is represented by the real points

* This convention is observed wherever double signs are used in the representation

$$\{-\alpha^2 + (2 \pm 3^{1/2})\alpha + 2 \pm 2(3^{1/2}), \quad \pm 3^{1/2}\alpha^2 + [3 \pm 2(3^{1/2})]\alpha, \quad -2(\alpha^2 + \alpha + 1)\},$$

which for varying α describe the circles

$$x^2 + y^2 \pm 3^{1/2}xz - yz - (1 \pm 3^{1/2})z^2 = 0.$$

The third pair of imaginary inflexions

$$(1 - \alpha\omega^2, \alpha\omega - 1, \omega^2 - \omega), \quad (1 - \alpha\omega, \alpha\omega^2 - 1, \omega - \omega^2)$$

is represented by the point-pair

$$[(1 \pm 3^{1/2})\alpha + 2, (1 \pm 3^{1/2})\alpha + 2, \mp 2(3^{1/2})],$$

whose locus as α varies is the line $x - y = 0$.

Thus it is found that *the Laguerre representation of the imaginary points of inflexion for the doubly infinite system of cubic curves with fixed real points of inflexion, two finite and one at infinity, and a fixed critic center, the one common to the three real harmonic polars, consists of two pairs of circles and a straight line counted twice.*

The straight line in the representation indicates a specialization due to having one fixed real inflexion at infinity. To obtain a more general result, the case when all three real inflexions are finite must be considered. It is here shown that the representatives of the three pairs of chains are not independent. They are related to each other and also to the fixed critic center, real harmonic polars and other associates of the cubic. These relations may be anticipated in the generalization.

4. *Chains of Critic Centers.* Before studying the more general form, the representatives of the critic centers for the system of cubics with one real inflexion at infinity may well be considered.

The fixed elements uniquely determine two imaginary critic centers, the Hessian points of I_1, I_2, I_3,

$$(\omega^2, 1, \omega), \quad (\omega, 1, \omega^2).$$

These are represented by two real points

$$H^{(1)}(1 + 3^{1/2}, 1 + 3^{1/2}, -2), \quad H^{(2)}(1 - 3^{1/2}, 1 - 3^{1/2}, -2)$$

on the line $x - y = 0$ where the four circles intersect in pairs.

The other three real critic centers are

$$V_1(\alpha + 1, -1, -1), \quad V_2(-1, \alpha + 1, -1), \quad V_3(-1, -1, \alpha + 1),$$

which as α varies describe the real harmonic polars h_1, h_2, h_3.

The remaining three pairs of imaginary critic centers are

$$(\alpha - 2\omega^2 + 1,\ \alpha\omega^2 - 1,\ \alpha\omega^2 - 1), \qquad (\alpha - 2\omega + 1,\ \alpha\omega - 1,\ \alpha\omega - 1);$$
$$(\alpha\omega^2 - 1,\ \alpha - 2\omega^2 + 1,\ \alpha\omega^2 - 1), \qquad (\alpha\omega - 1,\ \alpha - 2\omega + 1,\ \alpha\omega - 1);$$
$$(\alpha\omega^2 - 1,\ \alpha\omega^2 - 1,\ \alpha - 2\omega^2 + 1), \qquad (\alpha\omega - 1,\ \alpha\omega - 1,\ \alpha - 2\omega + 1).$$

These are represented by the real point-pairs

$$[-\alpha^2 - 7\alpha - 4,\ \alpha^2(2 \mp 3^{1/2}) + \alpha(2 \mp 3^{1/2}) + 2 \pm 2(3^{1/2}),\ 2(\alpha^2 + \alpha + 1)],$$
$$[\alpha^2(2 \pm 3^{1/2}) + \alpha(2 \pm 3^{1/2}) + 2 \mp 2(3^{1/2}),\ -\alpha^2 - 7\alpha - 4,\ 2(\alpha^2 + \alpha + 1)],$$
$$[\alpha^2(1 \pm 3^{1/2}) + \alpha(7 \pm 3^{1/2}) + 4 \mp 2(3^{1/2}),$$
$$\alpha^2(1 \mp 3^{1/2}) + \alpha(7 \mp 3^{1/2}) + 4 \pm 2(3^{1/2}),\ -2(\alpha^2 + 4\alpha + 7)],$$

which, independent of α, form the following three pairs of circles:

$$x^2 + y^2 + xz - (2 \pm 3^{1/2})yz - (1 \mp 3^{1/2})z^2 = 0,$$
$$x^2 + y^2 - (2 \mp 3^{1/2})xz + yz - (1 \pm 3^{1/2})z^2 = 0,$$
$$2(x^2 + y^2) - (1 \mp 3^{1/2})xz - (1 \pm 3^{1/2})yz - 2z^2 = 0.$$

These circles all pass through the point $\Sigma(1, 1, 1)$, with the second intersections of the mates of the pairs where the real harmonic polars meet the line through the real inflexions, that is at the Jacobian points of the three real inflexions. Further relations of these circles to each other and to the representations of the imaginary inflexions appear and hence may be anticipated in the generalization.

5. *Chains of Hessian Points.* For the further study let the three real points of inflexion be

$$I_1: (0, 1, -1), \quad I_2: (-1, 0, 1), \quad I_3: (a, -1, 1 - a),$$

where a may have any real value except zero and infinity, and I_3 is finite except when $a = 1$, which gives the special case that has been considered.

The Hessian points for I_1, I_2, I_3 are

$$(\omega^2 a, 1, -1 - \omega^2 a), \quad (\omega a, 1, -1 - \omega a),$$

represented by the real points

$$H^{(i)}: [2a^2 - a(1 \mp 3^{1/2}),\ -a(1 \mp 3^{1/2}) + 2,\ -2(a^2 - a + 1), \quad (i = 1, 2).$$

[illegible]

[illegible]

$$H_c^{(i)}: \pm 3^{1/2}(x^2 + y^2) + (1 \pm 3^{1/2})xz + (1 \pm 3^{1/2})yz + z^2 = 0, \quad (i = 1, 2).$$

These circles both pass through the points I_1 and I_2 and each passes through the center of the other.

Hence *the locus in a plane of the Laguerre representatives of the Hessian points with respect to two fixed real points and a collinear variable third real point is uniquely defined as two circles through the two fixed points such that each circle passes through the center of the other.*

As a takes all positive values from zero to infinity, the variable I_3 runs along the line segments exterior to I_1 and I_2; and the representatives of the corresponding Hessian points describe the outer arcs of the circles. As a takes all negative values from zero to infinity, I_3 runs along the interior segment from I_1 to I_2 and the representatives of the Hessians describe the inner arcs of the circles.

For I_3: $(a, -1, 1-a)$ the corresponding Hessian representatives are cut out on the circles by the line

$$x - y + (a^2 - 1)/(a^2 - a + 1) = 0.$$

The maximum and minimum values of $(a^2 - 1)/(a^2 - a + 1)$ are $\pm 2/3^{1/2}$; that is, $x - y \pm 2/3^{1/2} = 0$ are tangent lines to the circles, and each of the lines between the tangents cuts out two of the real point-pairs.

When $(a^2 - 1)/(a^2 - a + 1) = \pm 2/3^{1/2}$ the points I_3 are

$$M^{(1)}: (1 + 3^{1/2}, 1 - 3^{1/2}, -2), \quad M^{(2)}: (1 - 3^{1/2}, 1 + 3^{1/2}, -2).$$

As I_3 runs along the line segments exterior to these points, the representatives of the corresponding Hessian points describe the outer semicircles between the points of contact of the tangents. As I_3 runs along the line segment between these points the representatives of the Hessians describe the inner semicircles between the points of contact.

The Marie * representation of the imaginary points is obtained from the Laguerre by rotating each point-pair around the midpoint of the included line segment through an angle of ninety degrees. Thus imaginary points on a line are represented by pairs of real points on that line; and the Laguerre and Marie representatives of an imaginary point form the vertices of a square.

The points

$$M^{(1)}: (1 + 3^{1/2}, 1 - 3^{1/2}, -2), \quad M^{(2)}: (1 - 3^{1/2}, 1 + 3^{1/2}, -2)$$

* Marie, *Réalisation et usage des formes imaginaires en géométrie*, Paris (1891); Mouchot, *Les branches de la géométrie supérieure*, Paris (1892); Study, *Ausgewählte Gegenstände der Geometrie*, Leipzig (1911); Coolidge, *Geometry of the Complex Domain*, Oxford (1924).

which determine the Hessian points whose Laguerre representatives are the points of contact of the tangent lines are the Marie representatives of the Hessian points when $a = 1$, that is, when I_3 is the point at infinity.

It follows directly that the *locus in three-space of the Laguerre representatives of the Hessian points with respect to two fixed real points and a collinear variable third real point is a uniquely defined torus generated by revolving a circle about an axis which cuts it in two real points. As the variable third point runs along the segments of the axis exterior to the two fixed points, the representatives of the corresponding Hessian points describe the outer sheet of the surface. As the third point runs along the segment between the two fixed points, the Hessian representatives describe the inner sheet of the surface. When the third point has the positions of the Marie representatives of the Hessian points for the two fixed points and the point at infinity on the axis, the Laguerre representatives of the corresponding Hessian points describe the locus of parabolic points on the surface; and when the third point runs along the segments of the axis exterior to these Marie representatives, the Laguerre representatives of the corresponding Hessian points describe the synclastic part of the outer sheet of the surface.*

The Jacobian points with respect to

$$I_1:(0,1,-1),\quad I_2:(-1,0,1),\quad I_3:(a,-1,1-a)$$

are

$$J_1:(2a,-1,1-2a),\qquad J_2:(a/2,-1,1-a/2),\qquad J_3:(-a,-1,1+a).$$

It follows, because of the known mutual relations of these two triads of points, that the Hessian representatives for both the I and J points form the same two circles in the plane and the same torus in three-space.

6. *Circles as Representatives of Chains of Inflexions.* With

$$I_1:(0,1,-1),\quad I_2:(-1,0,1),\quad I_3:(a,-1,1-a)$$

as the real points of inflexion and $\Sigma(1,1,1)$ as the critic center common to the three real harmonic polars, the equations of the real harmonic polars are

$$h_1 \equiv \Sigma J_1 \equiv 3(x+2ay)-(1+2a)(x+y+z)=0,$$
$$h_2 \equiv \Sigma J_2 \equiv 3(2x+ay)-(2+a)(x+y+z)=0,$$
[illegible]

The Hessian lines of these harmonic polars are

$$3(x-\omega^i ay)-(1-\omega^i a)(x+y+z)=0,\qquad (i=1,2).$$

$$3x-(a+2+3\beta)(x+y+z)=0,$$
$$3ay-(1+2a+3\beta)(x+y+z)=0,$$
$$x+ay+\beta(x+y+z)=0.$$

The three pairs of imaginary inflexions are

$$(3\beta\omega^2a+\omega^2a^2+2\omega^2a,\quad 3\beta-\omega a+1,\quad -3\beta\omega^2a-3\beta-\omega^2a^2-a-1),$$
$$(3\beta\omega a+\omega a^2+2\omega a,\quad 3\beta-\omega^2a+1,\quad -3\beta\omega a-3\beta-\omega a^2-a-1);$$
$$(3\beta\omega^2a+\omega^2a^2-\omega a,\quad 3\beta+2a+1,\quad -3\beta\omega^2a-3\beta-\omega^2a^2-\omega^2a-1)$$
$$(3\beta\omega a+\omega a^2-\omega^2a,\quad 3\beta+2a+1,\quad -3\beta\omega a-3\beta-\omega a^2-\omega a-1);$$
$$(3\beta\omega^2a+\omega^2a^2-a,\quad 3\beta-\omega^2a+1,\quad -3\beta\omega^2a-3\beta-\omega^2a^2+2\omega a-1)$$
$$(3\beta\omega a+\omega a^2-a,\quad 3\beta-\omega a+1,\quad -3\beta\omega a-3\beta-\omega a^2+2\omega^2a-1)$$

The real representatives for the pairs of imaginary inflexions are

$$x=-9\beta^2a(2a-1\pm 3^{1/2})-3a[4a^2+2a-3\pm 2a(3^{1/2})]$$
$$-a[2a^3+3a^2-3a-2\pm(a^2-1)3^{1/2}],$$
$$y=9\beta^2[a-2\pm a(3^{1/2})]-3\beta[2a+4\pm a(2a+3)3^{1/2}]$$
$$-a^3-3a-2\mp a(a+3a+2)3^{1/2},$$
$$z=18\beta^2(a^2-a+1)+6\beta(2a^3-2a^2+a+2)+2(a^4-a^3+2a+1);$$

$$x=-9\beta^2a(2a-1\pm 3^{1/2})-3\beta a[4a^2+2a\pm(3a+2)3^{1/2}]$$
$$-a[2a^3+3a^2+1\pm(2a^2+3a+1)3^{1/2}],$$
$$y=9\beta^2[a-2-a(3^{1/2})]+3\beta[3a^2-2a-4-2a(3^{1/2})]$$
$$+2a^3+3a^2-3a-2\mp a(a^2-1)3^{1/2},$$
$$z=18\beta^2(a^2-a+1)+6\beta(2a^3+a^2-2a+2)+2(a^4+2a^3-a+1);$$

$$x=-9\beta^2a(2a-1\pm 3^{1/2})-3\beta a[4a^2+2a-3\pm(a+4)3^{1/2}]$$
$$-a[2a^3+3a^2-3a-2\pm 3(a+1)3^{1/2}],$$
$$y=9\beta^2[a-2\mp a(3^{1/2})]+3\beta[3a^2-2a-4\mp a(4a+1)3^{1/2}]$$
$$+2a^3+3a^2-3a-2\mp 3a^2(a+1)3^{1/2},$$
$$z=18\beta^2(a^2-a+1)-6\beta(2a^3+a^2+a+2)$$
$$+2(a^4+2a^3+3a^2+2a+1).$$

As β takes all values the representative point-pairs describe the following pairs of circles

$$2az(x-2y+z)+[\mp 3^{1/2}(x^2+y^2)$$
$$-(2\mp 3^{1/2})xz+yz+(1\pm 3^{1/2})z^2]=0,$$
$$2z(-2x+y+z)+a[\mp 3^{1/2}(x^2+y^2)+xz$$
$$-(2\mp 3^{1/2})yz+(1\pm 3^{1/2})z^2]=0,$$
$$\pm 3^{1/2}(x^2+y^2)-(2\pm 3^{1/2})xz+yz+(1\mp 3^{1/2})z^2$$
$$-a[\pm 3^{1/2}(x^2+y^2)+xz-(2\pm 3^{1/2})yz+(1\mp 3^{1/2})z^2]=0.$$

These circles all pass through Σ, the first of each pair through $H^{(1)}$ and the second of each pair through $H^{(2)}$.

The addition of the equations of the mates of the pairs of circles gives the equations of the real harmonic polars; and shows that the second real intersections of the mates are respectively on h_1, h_2, h_3.

Symbolically the equations of the three pairs of circles may be written as

$$\begin{aligned} 2az\Sigma I_2 + K_1 &= 0, & 2az\Sigma I_2 + K_2 &= 0; \\ 2z\Sigma I_1 + aK_3 &= 0, & 2z\Sigma I_1 + aK_4 &= 0; \\ K_2 - aK_4 &= 0, & K_1 - aK_3 &= 0. \end{aligned}$$

The lines $\Sigma I_2, \Sigma I_1$ cut the Hessian circles $H_c^{(1)}, H_c^{(2)}$, respectively, in pairs of points

$$\begin{aligned} &A^{(1)}[-3-(2)3^{\frac{1}{2}},\ 1-3^{\frac{1}{2}}, 5], && A^{(2)}[-3+2(3^{\frac{1}{2}}),\ 1+3^{\frac{1}{2}}, 5]; \\ &B^{(1)}[1-3^{\frac{1}{2}},\ -3-2(3^{\frac{1}{2}}), 5], && B^{(2)}[1+3^{\frac{1}{2}},\ -3+2(3^{\frac{1}{2}}), 5]; \end{aligned}$$

and the circles K_1, K_2, K_3, K_4 pass respectively through $A^{(1)}, A^{(2)}, B^{(1)}, B^{(2)}$. The circles K_2, K_4 intersect in Σ and

$$C^{(1)}[1-3^{\frac{1}{2}},\ 1-3^{\frac{1}{2}},\ +2(3^{\frac{1}{2}})];$$

the circles K_1, K_3 intersect in Σ and

$$C^{(2)}[1+3^{\frac{1}{2}},\ 1+3^{\frac{1}{2}},\ -2(3^{\frac{1}{2}})];$$

where $C^{(1)}$ and $C^{(2)}$ are the centers of the circles $H_c^{(1)}, H_c^{(2)}$.

Hence *the Laguerre representation of the imaginary points of inflexion for the doubly infinite system of plane cubic curves with three real points of inflexion I_1, I_2, I_3 and a fixed critic center Σ, the one com'mon to the three real harmonic polars h_1, h_2, h_3, consists of three uniquely defined pairs of circles*

$$\Sigma H^{(1)} A^{(1)},\quad \Sigma H^{(2)} A^{(2)};\quad \Sigma H^{(1)} B^{(1)},\quad \Sigma H^{(2)} B^{(2)};\quad \Sigma H^{(1)} C^{(1)},\quad \Sigma H^{(2)} C^{(2)};$$

and the mates of these pairs have their second real intersections on h_1, h_2, h_3, respectively.

7. *Pencils of Representative Circles.* The form;

$$\begin{aligned} 2az\Sigma I_2 + K_1 &= 0, & 2az\Sigma I_2 + K_2 &= 0; \\ 2z\Sigma I_1 + aK_3 &= 0, & 2z\Sigma I_1 + aK_4 &= 0; \\ K_2 - aK_4 &= 0, & K_1 - aK_3 &= 0; \end{aligned}$$

Hence the Laguerre representatives of the imaginary inflexions as I_3 runs along the line $I_1 I_2$ are pairs of pencils of circles on the points

$$\Sigma, A^{(1)} \text{ and } \Sigma, A^{(2)}; \quad \Sigma, B^{(1)} \text{ and } \Sigma, B^{(2)}; \quad \Sigma, C^{(1)} \text{ and } \Sigma, C^{(2)}.$$

8. *Pencils of Representative Critic Centers.* Methods similar to those used for the inflexions give that the imaginary critic centers for the cubics considered, in addition to the Hessian points, are represented by the pairs of pencils of circles

$$\begin{aligned}
&\mp 3^{1/2}(x^2+y^2)-(2\mp 3^{1/2})xz+yz+(1\pm 3^{1/2})z^2\\
&\qquad +2a[\pm 3^{1/2}(x^2+y^2)+xz-(2\pm 3^{1/2})yz+(1\mp 3^{1/2})z^2]=0,\\
&\mp 3^{1/2}(x^2+y^2)-(2\mp 3^{1/2})xz+yz-(1\pm 3^{1/2})z^2\\
&\qquad +(a/2)[\pm 3^{1/2}(x^2+y^2)+xz-(2\pm 3^{1/2})yz+(1\mp 3^{1/2})z^2]=0,\\
&\mp 3^{1/2}(x^2+y^2)-(2\mp 3^{1/2})xz+yz-(1\pm 3^{1/2})z^2\\
&\qquad -a[\pm 3^{1/2}(x^2+y^2)+xz-(2\pm 3^{1/2})yz+(1\mp 3^{1/2})z^2]=0;
\end{aligned}$$

that is,

$$\begin{aligned}
&K_1-2aK_4=0, &&K_2-2aK_3=0;\\
&K_1-(a/2)K_4=0, &&K_2-(a/2)K_3=0;\\
&K_1-aK_4=0, &&K_2-aK_3=0.
\end{aligned}$$

These circles are all through Σ; the first circle of each pair is through $M^{(1)}$ and the second of each pair through $M^{(2)}$, where $M^{(1)}$ and $M^{(2)}$ are the Marie representatives of the Hessian points when I_3 is at infinity; and the mates of the pairs intersect respectively in J_1, J_2, J_3.

Hence the Laguerre representatives of the imaginary critic centers, as I_3 runs along the line $I_1 I_2$, are the two Hessian circles $H_c^{(1)}, H_c^{(2)}$ and pencils of circles on

$$\Sigma, M^{(1)} \text{ and } \Sigma, M^{(2)},$$

with the second real intersections of the mates of the pairs at the Jacobian points of I_1, I_2, I_3.

WEST VIRGINIA UNIVERSITY.

A Character Symbol for Primes Relative to a Cubic Field.

By Chester G. Jaeger.

1. *Introduction.* Any cubic field defined by a binomial equation may have the defining equation written in the form

$$x^3 - ab^2 = 0$$

where a and b are relatively prime, and neither has a perfect square of a rational prime as a factor. In his study of the pure cubic field Dedekind * separates the types of fields into two kinds, I and II, according to whether $a^2 - b^2$ is not or is congruent to zero, mod 9. For I let $k = ab$, and for II let $k = 3ab$.

It has been shown that the class number, h, of a field is determined by the equation

$$gh = \lim_{s \to 1} (s - 1) \zeta_K(s)$$

where g is a constant dependent on the field, and $\zeta_K(s)$ is Dedekind's zeta function for a field.

In order to determine the class number in the pure cubic field, Dedekind studied the factorization of rational primes in the quadratic field $K(\sqrt{-3})$; and from this he defined a function $\Psi(\mathfrak{p})$ which is the character of the prime ideals, as follows:

If $\mathfrak{p}$ is a factor of k, $\Psi(\mathfrak{p}) = 0$.
If $\mathfrak{p}$ is a factor of 3, but not of k, $\Psi(\mathfrak{p}) = 1$.
For all other prime ideals, $\Psi(\mathfrak{p}) = (ab^2/\mathfrak{p})_3$,

where $(ab^2/\mathfrak{p})_3 = \rho^i$, $(i = 1, 2, 3)$, and where ρ is a primitive cube root of unity. This symbol is the cubic residue character of ab^2 with respect to the modulus $\mathfrak{p}$. He then shows that we may write

$$\zeta_K(s) = \prod_p F(p),$$

where the product extends over all the rational primes, and

$$F(p) = [1/(1 - 1/p^s)] \cdot \prod 1/[1 - \Psi(\mathfrak{p})/N(\mathfrak{p})^s],$$

the product extending over the prime ideal factors of p.

It is the object of this paper to define a function similar to $\Psi(\mathfrak{p})$ for a general cubic field. This will be accomplished in the following steps:

(1) It will be determined what values Ψ must have when the various types of primes, p, are factored in a certain quadratic field.

(2) The general cubic equation, $F_3(x) = 0$, will be transformed into a binomial cubic, $y^3 - \beta = 0$.

(3) Letting Δ_3 be the discriminant of $F_3(x) = 0$, and Δ_2 the discriminant of $F_2(x) = 0$, a quadratic equation related to $F_3(x) = 0$, the cubic character of β will be determined in the fields, $K(\Delta_2^{1/2})$ and $K(\omega, \Delta_2^{1/2})$. Here ω is a primitive cube root of unity.

(4) β is a cubic residue, mod p, if $y^3 - \beta$ has a solution in $K(p, \Delta_2^{1/2})$. This property is indicated by $\{\beta/\pi\} = \omega^i$, $i = 0, 1, 2$, depending on the type of prime. π is a factor of p in $K(\Delta_3^{1/2})$.

(5) If $\Psi(\pi) = \{\beta/\pi\}$, then, it will be shown that

$$\zeta_K(s) = \prod_p [1/(1 - 1/p^s)] \cdot \prod 1/[1 - \Psi(\mathfrak{p})/N(\mathfrak{p})],$$

the first product extending over all rational primes, and the second product over all the prime ideals of $K(\Delta_3^{1/2})$.

2. *Types of Primes.* Every rational prime p defines a principal ideal. The various ways in which this ideal may be resolved into its prime ideal factors in a cubic number field plays an important part in this article. A prime may be any one of five types—p_1, p_2, p_3, p_4, or p_5—according to the kinds of factors it has. Following is a table showing this, together with the norms of the factors.

Type	*Factors*	*Norms*
p_1	$\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$	$N(\mathfrak{p}_1) = N(\mathfrak{p}_2) = N(\mathfrak{p}_3) = p_1$
p_2	$\mathfrak{p}_1\mathfrak{p}_2$	$N(\mathfrak{p}_1) = p_2$, $N(\mathfrak{p}_2) = p_2^2$
p_3	$\mathfrak{p}_1$	$N(\mathfrak{p}_1) = p_3^3$
p_4	$\mathfrak{p}_1\mathfrak{p}_2^2$	$N(\mathfrak{p}_1) = N(\mathfrak{p}_2) = p_4$
p_5	$\mathfrak{p}_1^3$	$N(\mathfrak{p}_1) = p_5$.

The factors of the primes of the types p_1, p_4, and p_5 are of the first degree. For the type p_2, $\mathfrak{p}_1$ is of the first degree, and $\mathfrak{p}_2$ is of degree 2. For the type p_3, $\mathfrak{p}_1$ is of degree 3. The notation, p_i, just introduced shall be adhered to throughout this article so that the subscript shall indicate the type of rational prime.

3. *The Function* $\Psi(\pi)$. In a quadratic field there are only three types

of factorization of a rational prime, p. That is, p is the product of two distinct primes, $p = \pi' \cdot \pi''$; or p is the square of a prime, $p = \pi^2$; or p remains prime, $p = \pi$. We shall associate with the general cubic field a certain quadratic field, Q, and then define a function, $\Psi(\pi)$, of all the prime ideals in Q such that

$$(1) \qquad \lim_{s \to 1} (s - 1)\zeta_K(s) = \prod_{\pi} 1/[1 - \Psi(\pi)/N(\pi)],$$

π ranging over all the prime ideals of Q.

For any field, $\zeta_K(s)$ is made up of the product of all the fractions of the type

$$1/[1 - 1/N(\mathfrak{p})^s]$$

where $\mathfrak{p}$ ranges over all the prime ideals of the field. By grouping those ideals which arise from each of the types of the rational primes, p_i, and using the tabulated values of the norms in each case together with the fact that $(s - 1)$ times the Riemann zeta function approaches unity as s approaches 1, we may write, for the cubic field,

$$(2) \qquad \begin{aligned} &\lim_{s \to 1} (s - 1)\zeta_K(s) \\ &\quad = \prod_{p_1} [1/(1 - 1/p_1)]^2 \cdot \prod_{p_2} [1/(1 - 1/p_2^2)] \\ &\qquad \times \prod_{p_3} [1/(1 + 1/p_3 + 1/p_3^2)] \cdot \prod_{p_4} [1/(1 - 1/p_4)]. \end{aligned}$$

Thus, we see, by referring to (1) that Q must be so chosen, and $\Psi(\pi)$ so defined that the following must hold. (Here the norms are in Q).

For a prime of the type p_1,

$N(\pi') = N(\pi'') = p_1$, in which case $p_1 = \pi' \cdot \pi''$ in Q and $\Psi(\pi') = \Psi(\pi'') = 1$.

For a prime of the type p_2,

$N(\pi) = p_2^2$, in which case $p_2 = \pi$ in Q, and $\Psi(\pi) = 1$; or,

$N(\pi') = N(\pi'') = p_2$, in which case $p_2 = \pi' \cdot \pi''$ in Q, and $\Psi(\pi') = -\Psi(\pi'') = \pm 1$.

(We shall see later, § 5, that this second factorization of [illegible]).

For [illegible] of the type [illegible]

[illegible] in which [illegible] $\pi' \cdot \pi''$ in [illegible]

For a prime of the type p_4,

$N(\pi) = p_4$, in which case $p_4 = \pi^2$ in Q, and $\Psi(\pi) = 1$.

For a prime of the type p_5,

$N(\pi) = p_5$, or $N(\pi') = N(\pi'') = p_5$.

In either case $\Psi(\pi) = 0$, or $\Psi(\pi') = \Psi(\pi'') = 0$.

4. *The General Cubic Equation whose Roots Determine the Cubic Field.* There is no loss in generality in considering the general cubic equation to be of the form

$$x^3 + Cx + D = 0$$

since by linear transformation every cubic can be so expressed. Let us call this cubic $F_3(x)$. Now if C contains a power of p greater than or equal to 2, and if D contains a power of p greater than or equal to 3, then the roots of $F_3(x) = 0$ could be divided by a power of p, thus making a further reduction of the equation. We shall therefore assume that either p occurs to a power less than 2 in C or to a power less than 3 in D.

Closely associated with $F_3(x)$ is the quadratic

$$F_2(x) = 3Cx^2 + 9Dx - C^2 = 0.$$

If we let Δ_3 and Δ_2 be, respectively, the discriminants of $F_3(x) = 0$ and $F_2(x) = 0$, then it is easily verified that

$$\Delta_2 = -3\Delta_3.$$

Call the roots of $F_2(x) = 0$, μ_1 and μ_2. Then, if we apply the transformation

$$x = (\mu_1 y + 3\mu_1\mu_2)/(y + 3\mu_1) \tag{3}$$

to $F_3(x) = 0$, it becomes the binomial cubic

$$y^3 - 9C\mu_1 = 0.$$

And, finally, for convenience of notation, if we let $\beta = 9C\mu_1$, we have

$$y^3 - \beta = 0, \tag{4}$$

which has its coefficients in $K(\Delta_2^{1/2})$.

5. *The Quadratic Character of* Δ_3, mod p. In the field $K(p_1)$, $F_3(x)$ has three linear factors.

$$F_3(x) = (x - a_1)(x - a_2)(x - a_3) \qquad (p_1)$$

Obviously,

$$\Delta_3 = (a_1 - a_2)^2(a_1 - a_3)^2(a_2 - a_3)^2 \qquad (p_1)$$

is a perfect square. From this it follows that $(\Delta_3/p_1) = 1$. Hence, in $K(\Delta_3^{1/2})$, $p_1 = \pi' \cdot \pi''$.

In the field $K(p_2)$, $F_3(x)$ is factorable into a linear and a quadratic factor,

$$F_3(x) = (x - a)(x^2 - bx + c) \qquad (p_2).$$

In this case

$$\begin{aligned}\Delta_3 &= D(x-a)D(x^2 - bx + c)R[(x-a),\ (x^2 - bx + c)]^2 \\ &= D(x^2 - bx + c)R[(x-a),\ (x^2 - bx + c)]^2.\end{aligned}$$

Hence, we see that the discriminant of $F_3(x)$ differs from the discriminant of $x^2 - bx + c$ only by a square factor. But this means that the quadratic field $K(\Delta_3^{1/2})$ must be identical with the quadratic field determined by the roots of the equation

$$x^2 - bx + c = 0. \qquad (p_2)$$

This last equation is irreducible in $K(p_2)$ and hence its discriminant could not be a quadratic residue module p_2.* Therefore $(\Delta_3/p_2) = -1$, or, in $K(\Delta_3^{1/2})$, $p_2 = \pi$. (Note that this excludes one of the possibilities of factoring p_2 mentioned in § 3).

In $K(p_3)$, $F_3(x)$ is irreducible. And for $p_3 > 3$, p_3 is relatively prime to Δ_3.† Let α be a number satisfying the congruence

$$F_3(x) \equiv 0, \bmod p_3.$$

Then, since p_3 is relatively prime to Δ_3, $F_3(x)$ is a prime function, mod p_3. Therefore,

$$F_3(x) \equiv (x - \alpha)(x - \alpha^{p_3})(x - \alpha^{p_3^2}), \bmod p_3; \text{ or}$$
$$\Delta_3 \equiv \delta^2, \bmod p_3, \text{ where } \delta = (\alpha - \alpha^{p_3})(\alpha - \alpha^{p_3^2})(\alpha^{p_3} - \alpha^{p_3^2}).$$

δ is rational, mod p_3, since it is unchanged by the substitution, (α, α^{p_3}). Hence Δ_3 is a quadratic residue, mod p_3. So, $(\Delta_3/p_3) = 1$; and in $K(\Delta_3^{1/2})$, $p_3 = \pi' \cdot \pi''$.

We shall now apply the theorem of Dedekind to

$$p_1 = \mathfrak{p}_1 \mathfrak{p}_2^2.$$

Since we are considering only $p_1 > 3$, the discriminant of $K(\alpha)$‡ is divisible by p_1 with the exponent 1. It is a well known fact that the discriminant of

* Wahlin, *American Journal of Mathematics*, Vol. 41. If p_2 is contained in Δ_3 it is to an even power, and hence $(\Delta_3/p_2) = (\Delta_3'/p_2)$, where $\Delta_3 = p_2^a \Delta_3'$.

an equation differs from the field discriminant by a square factor. Therefore Δ_3 is divisible by p_4 an odd number of times, since this square factor, if it contains p_4 at all must contain it an even number of times. Now in $K(p_4)$, $F_3(x)$ has a quadratic and a linear factor,

$$F(x) = (x - a)(x^2 - bx + c) \qquad (p_4).$$

As in the case of p_2, we could show that

$$\Delta_3 = A^2 D(x^2 - bx + c).$$

Now since Δ_3 contains p_4 an odd number of times, $D(x^2 - bx + c)$ must contain it an odd number of times, and therefore at least once. Thus the discriminant of $K(\Delta_3^{1/2})$ is divisible by p_4. So, in $K(\Delta_3^{1/2})$, $p_4 = \pi^2$.

6. *The Cubic Character of β in $K(\Delta_2^{1/2})$ with respect to π_1 and π_2.* If

$$F_3(x) = (x - r_1)(x - r_2)(x - r_3), \qquad (p)$$

then by applying the transformation (3) we have

$$(5) \qquad y^3 - \beta = (y - \rho_1)(y - \rho_2)(y - \rho_3) \qquad (p)$$

where ρ_1, ρ_2 and ρ_3 are numbers in the Ring $R(p, \Delta_2^{1/2})$.

And, if

$$F_3(x) = (x - r)(x^2 + mx + n) \qquad (p)$$

then the transformation gives

$$(6) \qquad y^3 - \beta = (y - \rho)(y^2 + \mu y + \nu) \qquad (p)$$

where ρ, μ and ν are numbers in the Ring $R(p, \Delta_2^{1/2})$.

Equations (5) and (6) are possible when and only when β is a cubic residue mod p in the field $K(\Delta_2^{1/2})$. And by Euler's Criterion this is true when and only when

$$\beta^{[N(\pi_1)-1]/\delta} \equiv 1, \bmod \pi_1,$$

where δ is the greatest common division of $N(\pi_1) - 1$ and 3, and where π_1 is any prime factor of p in $K(\Delta_2^{1/2})$. Now β may contain π_1,

$$\{\beta\} = \pi_1^k \bar{\bar{\beta}}$$

$\bar{\bar{\beta}}$ being prime to π_1.

If $k \equiv 0$ mod 3 and π_1 is a principal ideal, then dividing the roots of (4) by $\pi_1^{k/3}$ would give

$$y^3 - \bar{\beta} = 0.$$

If $k \equiv 0 \bmod 3$ and π_1 is not a principal ideal, then the principal ideal $\{\beta\}$ is also divisible by π_1^k. Now there exists an ideal I, such that I and $\{\beta\}$ are relatively prime, and such that $I \cdot \pi_1$, is a principal ideal, $\{\gamma\}$. Since I and $\{\beta\}$ are relatively prime there is an integer in I which is prime to β. Call this integer λ. Now, if we multiply the roots of (4) by $\lambda^{k/3}$ and divide them by $\gamma^{k/3}$ we get

$$y^3 - - \bar{\beta} = 0.$$

If $k \not\equiv 0 \bmod 3$, then it can be shown that equation (4) is irreducible, and that further π_1 is the cube of a prime ideal.*

Let β' be the conjugate of β. A short calculation shows that

$$\beta\beta' = -27C^3.$$

Thus, if β is a cubic residue mod π_1, β' is also a cubic residue mod π_1; and β' is a cubic residue with respect to the conjugate ideal π_2. Hence, from the relation existing between β and β', we can conclude that if β' is a cubic residue mod π_1, it is also a cubic residue mod π_2. (Of course, if p remains prime in $K(\Delta_2^{1/2})$, then this last consideration is superfluous).

7. *The Cubic Character of β in the Field $K(\omega, \Delta_2^{1/2})$ with Respect to π' and π''.* Since $\Delta_2 = -3\Delta_3$, we see immediately that the composite field $K(\Delta_3^{1/2}, \Delta_2^{1/2})$ is the same as the field $K(\omega, \Delta_2^{1/2})$. We shall show that β is a cubic residue mod π' and π'' if it is a cubic residue mod π_1 and π_2. This will depend on the ways in which p is factorable in the two fields $K(\Delta_2^{1/2})$ and $K(\Delta_3^{1/2})$. Consider the four possibilities (A), (B), (C), and (D):

(A) In $K(\Delta_2^{1/2})$, $p = \pi_1 \cdot \pi_2$
In $K(\Delta_3^{1/2})$, $p = \pi' \cdot \pi''$.

We saw that in this case (§ 5) Δ_3 is a quadratic residue in the rational field mod p, and hence certainly mod π_1 and mod π_2 in the field $K(\Delta_2^{1/2})$. Thus, in the field $K(\omega, \Delta_2^{1/2})$, π_1 and π_2 are each factorable into two distinct factors. $\pi_1 = \pi_1' \cdot \pi_1''$, $\pi_2 = \pi_2' \cdot \pi_2''$; and from unique factorization $\pi' = \pi_1'\pi_2'$ and $\pi'' = \pi_1'' \cdot \pi_2''$ in the field $K(\omega, \Delta_2^{1/2})$. Thus, if β is a cubic residue mod π_1 and π_2 in $K(\omega, \Delta_2^{1/2})$, then it is certainly a cubic residue mod π_1', π_2'. [illegible] mod π' and mod π'' [illegible] $K(\omega, \Delta_2^{1/2})$. We shall indicate this cubic residue property by

$$\{\beta/\pi'\} = \{\beta/\pi''\} = +1.$$

(B) In $K(\Delta_2^{1/2})$, $p = \pi_1 \cdot \pi_2$
In $K(\Delta_3^{1/2})$, $p = \pi'$.

In this case, we saw (§ 5) that Δ_3 is not a quadratic residue mod $\pi' = p$. And since $\Delta_2 = -3\Delta_3$ is a quadratic residue, -3 must be a quadratic non-residue mod p. Hence

$$p + 1 \equiv 0, \text{ mod } 3.$$

In this case, then, $\delta = 1$. By hypothesis

$$\beta^{p-1} \equiv 1, \text{ mod } \pi_1 \text{ and } \pi_2.$$

Then surely

$$\beta^{p-1} \equiv 1, \text{ mod } \pi'.$$

$(p + 1)/3$ is an integer. Raising both sides to the power $(p + 1)/3$ we have

$$\beta^{(p^2-1)/3} \equiv 1, \text{ mod } \pi'.$$

Now, since $N(\pi')$* $= p^2$, this becomes

$$\beta^{[N(\pi')-1]/3} \equiv 1, \text{ mod } \pi' \text{ or}$$
$$\{\beta/\pi'\} = 1$$

(C) In $K(\Delta_2^{1/2})$, $p = \pi_1$
In $K(\Delta_3^{1/2})$, $p = \pi' \cdot \pi''$.

We see at once that $N(\pi_1) = p^2$. And since

$$\beta^{[N(\pi_1)-1]/3} = \beta^{(p^2-1)/3} \equiv 1, \text{ mod } \pi_1,$$

it follows that

$$\beta^{(p^2-1)/3} \equiv 1, \text{ mod } \pi' \text{ and } \pi''.$$

This says that in $K(\omega, \Delta_2^{1/2})$, β is cubic residue mod π' and π''. Or,

$$\{\beta/\pi'\} = \{\beta/\pi''\} = 1.$$

(D) In $K(\Delta_2^{1/2})$, $p = \pi_1$,
In $K(\Delta_3^{1/2})$, $p = \pi'$.

Here $(\Delta_3/p) = -1$, and $(\Delta_2/p) = -1$. Thus $(\Delta_2\Delta_3/p) = +1 = (-3\Delta_3^2/p) = (-3/p)$. So in $K(\omega, \Delta_2^{1/2})$, p may have two factors $\pi^{(1)}$, $\pi^{(2)}$. Then since $\beta^{(p^2-1)/3} \equiv 1$, mod π_1, it follows that in the field $K(\omega, \Delta_2^{1/2})$, $\beta^{(p^2-1)/3} \equiv 1$, mod $\pi^{(1)}$ and $\pi^{(2)}$, and hence, certainly

$$\beta^{(p^2-1)/3} \equiv 1, \text{ mod } \pi'.$$

Or, we may write

---

* It is understood in all cases, norm is taken in the field $K(\Delta_3^{1/2})$.

$$\beta^{[N(\pi)-1]/3} \equiv 1, \bmod \pi'.$$

Thus, for this case also we have

$$\{\beta/\pi\} = 1.$$

8. *The Function* $\Psi(\pi)$. If in $K(\Delta_2^{1/2})$ $p = \pi_1 \cdot \pi_2$ and β is a cubic residue mod π_1 and mod π_2 it follows that it is a cubic residue mod p, provided p does not happen to be a factor of Δ_2,

$$\Delta_2 = p^s \Delta_2'$$

with s odd. For p larger than 3, p would be contained in Δ_3 to the same power as in Δ_2. From the theorem of Dedekind we see that primes of the types p_1 and p_2 are not contained in the field discriminant of $K(\alpha)$, and hence if in Δ_3, it must be to an even power.

We also saw that in $K(\Delta_3^{1/2})$, $p_1 = \pi' \cdot \pi''$. So this fits the case (A) just considered. That is, for primes of the type p_1,

$$\{\beta/\pi'\} = \{\beta/\pi''\} = 1.$$

Let us put $\Psi(\pi) = \{\beta/\pi\}$. Then, since $N(\pi') = N(\pi'') = p_1$, we have

$$1/(1-1/p_1)^2 = \prod_{i=1}^{2} 1/[1-\Psi(\pi^{(i)})/N(\pi^{i})].$$

And in $K(\Delta_3^{1/2})$, $p_2 = \pi'$. This is the case, (B) or (D), where $\{\beta/\pi'\} = 1$. Here, $N(\pi') = p_2^2$. So

$$1/(1-1/p_2^2) = 1[1-\Psi(\pi')/N(\pi')].$$

In § 6, equations (5) and (6) are the cases which arose from $F_3(x)$ being factorable in the ring $R(p)$. Let us next take up the case where $F_3(x)$ is irreducible in the ring $R(p)$. This is the case where p is of the type p_3. In this case $F_3(x)$ can not be resolved into factors involving only quadratic irrationalities. Thus $y^3 - \beta$ must also be irreducible in the ring $R(p, \Delta_2^{1/2})$, and β could not be a cubic residue. For, if $y^3 - \beta$ were factorable in $R(p, \Delta_2^{1/2})$, $y^3 - \beta = 0$ would have at least one root in this ring. Then $F_3(x)$ would have a root in $R(p, \Delta_2^{1/2})$. But, since $F_3(x)$ is irreducible in $R(p)$, it cannot have factors involving only quadratic irrationalities.

Thus for primes of the type p_3, β is not a cubic residue. In order to find the function $\Psi(\pi)$ we have two cases to consider: (a) when $p_3 = \pi_1 \cdot \pi_2$ in $K(\Delta_2^{1/2})$ and (b) when $p_3 = \pi_1$ in $K(\Delta_2^{1/2})$. In either case, we have seen, page ——, that $p_3 = \pi' \cdot \pi''$ in $K(\Delta_3^{1/2})$.

(a) This corresponds to (A) above. In this case we saw that $\pi_1 = \pi_1'\pi_1''$

$$\beta^{p_3-1}-1\equiv 0, \text{ mod } \pi_1 \text{ and } \pi_2$$

since π_1 and π_2 are primes of the first degree in $K(\Delta_2^{1/2})$. Moreover,

$$\beta^{(p_3-1)/3}\not\equiv 1, \text{ mod } \pi_1 \text{ or } \pi_2$$

since β is not a cubic residue mod π_1 or π_2. Thus, if we write

$$\beta^{p_3-1}-1=[\beta^{(p_3-1)/3}-1]\,[\beta^{(p_3-1)/3}-\omega]\,[\beta^{(p_3-1)/3}-\omega^2]\equiv 0, \text{ mod } \pi_1 \text{ and } \pi_2$$

it is obvious that

$$[\beta^{(p_3-1)/3}-\omega]\,[\beta^{(p_3-1)/3}-\omega^2]\equiv 0, \text{ mod } \pi_1 \text{ and } \pi_2.$$

Therefore, $\beta^{(p_3-1)/3}-\omega$ is divisible by a factor of π_1 or $\beta^{(p_3-1)/3}-\omega^2$ is divisible by π_1. But, since $\beta^{(p_3-1)/3}$ is congruent to a rational integer, b, mod π_1, if $\beta^{(p_3-1)/3}-\omega^2$ is divisible by π_1, ω^2-b is divisible by π_1. However, ω^2 is the root of an irreducible equation of degree 2 relative to $K(\Delta_2^{1/2})$, and π_1 is a prime ideal in this field which is not a divisor of the index of ω^2. Hence, ω^2 cannot satisfy a congruence of degree less than 2 mod π_1, and thus

$$\beta^{(p_3-1)/3}-\omega^2\equiv 0, \text{ mod } \pi_1$$

is impossible.

In the same way, $\beta^{(p_3-1)/3}-\omega$ is not divisible by π_1. But, $[\beta^{(p_3-1)/3}-\omega]\,[\beta^{(p_3-1)/3}-\omega^2]$ is divisible by π_1 and by π_2 and hence by the four factors π_1', π_2', π_1'' and π_2''. Two of these must be divisors of $\beta^{(p_3-1)/3}-\omega$ and two of $\beta^{(p_3-1)/3}-\omega^2$. Let us assume that the notation is such that π_1' is a divisor of $\beta^{(p_3-1)/3}-\omega$. Then the other divisor is π_2' or π_2''. First, assume it is π_2''. Then $\beta^{(p_3-1)/3}-\omega^2$ is divisible by π_2', and β is congruent to a rational integer, c, mod π_2'; it follows that

$$c-\omega^2\equiv 0, \text{ mod } \pi_2'.$$

Considering $K(\omega, \Delta_2^{1/2})$ as a relative field to $K(\Delta_3^{1/2})$, it follows from a change of $-\Delta_2^{1/2}$ into $\Delta_2^{1/2}$ that $(-3)^{1/2}=\Delta_2^{1/2}/\Delta_3^{1/2}$ is transformed into $-(-3)^{1/2}$. Consequently, from $b-\omega\equiv 0$, mod π_1', follows $b-\omega^2\equiv 0$, mod π_2'. Therefore

$$b\equiv c, \text{ mod } p_3.$$

Let β' be the conjugate of β. Then, from

$$b\equiv\beta^{(p_3-1)/3}, \text{ mod } \pi_1$$

we have

(7) $$b\equiv\beta'^{(p_3-1)/3}, \text{ mod } \pi_2;$$

also

(8) $$\beta^{(p_3-1)/3}\equiv c\equiv b, \text{ mod } \pi_2.$$

Thus, from (7) and (8) it follows that

(9) $$(\beta\beta')^{(p_3-1)/3}\equiv b^2, \text{ mod } \pi_2.$$

But $\beta\beta' = N(\beta) = -27C^3$, and hence

$$(\beta\beta')^{(p_3-1)/3} \equiv (-3C)^{p_3-1} \equiv 1, \bmod \pi_2.$$

Then, from (8) and (9)

$$\beta^{2[(p_3-1)/3]} \equiv 1, \bmod \pi_2 \quad \text{or} \quad \beta^{(p_3-1)/3} \equiv 1, \bmod \pi_2,$$

which contradicts the assumption that $\beta^{(p_3-1)/3} \not\equiv 1$, mod π_2. Therefore, $\beta^{(p_3-1)/3} - \omega \equiv 0$, mod π_2'' is impossible. We must thus conclude that

$$\beta^{(p_3-1)/3} \equiv 0, \bmod \pi_2'.$$

And finally, therefore,

$$\text{(10)} \qquad \beta^{(p_3-1)/3} - \omega \equiv 0, \bmod \pi_1'\pi_2' (= \pi').$$

If p_3 is a divisor of the index of $\beta^{(p_3-1)/3}$ then

$$\beta^{(p_3-1)/3} \equiv \beta'^{(p_3-1)/3}, \bmod p_3.$$

But then

$$\beta^{2[(p_3-1)/3]} \equiv (\beta\beta')^{(p_3-1)/3} \equiv 1, \bmod p_3 \quad \text{or} \quad \beta^{(p_3-1)/3} \equiv 1, \bmod p_3,$$

contrary to our assumption. Hence p_3 cannot be a divisor of the index of $\beta^{(p_3-1)/3}$.

So, for $p_3 = \pi_1 \cdot \pi_2$ in $K(\Delta_2^{1/2})$,

$$\Psi(\pi') = \omega, \quad \text{and} \quad \Psi(\pi'') = \omega^2.$$

And, since $N(\pi') = N(\pi'') = p_3$

$$\text{(11)} \quad [1/(1-\omega/p_3)]\,[1/(1-\omega^2/p_3)] = \prod_{i=1}^{2} \{1/[1-\Psi(\pi^{(i)})/N(\pi^{(i)})]\}.$$

(b) If in $K(\Delta_2^{1/2})$ p_3 remains prime, then by Fermat's theorem

$$\beta^{p_3^2-1} - 1 \equiv 0, \bmod \pi_1,$$

since then π_1 is a prime of degree 2 in $K(\Delta_2^{1/2})$.

In a manner precisely similar to the case just treated, it is easily seen that

$$\beta^{(p_3^2-1)/3} \equiv \omega, \bmod \pi' \quad \text{and} \quad \beta^{(p_3^2-1)/3} \equiv \omega^2, \bmod \pi''.$$

Thus $\Psi(\pi') = \omega$, $\Psi(\pi'') = \omega^2$ and, in either case (a) or (b), for a prime of the type p_3, the equation (11) holds. Since for p_4

$$F_3(x) = (x-a)(x^2 - bx + c), \qquad (p_4)$$
$$y^3 - \beta = (y-\rho)(y^2 + \mu y + \nu) \qquad (p_4)$$

in $K(\rho, \Delta_2^{1/2})$ and hence β is a cubic residue.

$$\Psi(\pi') = \{\beta/\pi'\} = 1, \quad N(\pi') = p_4, \quad 1/(1-1/p_4) = 1/[1-\Psi(\pi)/N(\pi)].$$

9. $\Psi(\pi)$ *for* 3 *and* 2. Nothing in this discussion has excluded the primes 3 and 2 except when they are of the types p_3 and p_4. Wahlin has shown* that when 3 is of the type p_3, then either $s=0$, and $\Delta_3 \equiv 1$, mod 3 or $s=6$ and $\Delta_3' \equiv 1$, mod 3. In either case, Δ_3 is a quadratic residue mod 3, and hence in $K(\Delta_3^{1/2})$, $3=\pi' \cdot \pi''$.
So we shall define

$$\{\beta/\pi'\}=\omega, \qquad \{\beta/\pi''\}=\omega^2.$$

The only case where 2 remains prime in a cubic field is when Δ_3 is odd and $C \not\equiv 0$ mod 2. Now $\Delta_3 = -27D^2-4C^3$. The following congruence is then obvious:

$$\Delta_3 = -27D^2-4C^3 \equiv 1, \text{ mod } 8;$$

and hence, $2=\pi' \cdot \pi''$; again we shall define

$$\{\beta/\pi'\}=\omega, \qquad \{\beta/\pi''\}=\omega^2.$$

When 3 is of the type p_4, from this same article by Wahlin, it can be shown that in all cases s is odd, and therefore $3=\pi^2$. So, we shall, since $y^3-\beta$ is reducible in $R(2, \Delta_2^{1/2})$, define

$$\{\beta/\pi\}=1.$$

When 2 is of the type p_4, s is either odd, in which case $2=\pi^2$; or else s is even in which case $2=\pi'\pi''$. In either case, we shall set

$$\Psi(\pi')=1 \quad \text{or} \quad \Psi(\pi')=\Psi(\pi'')=1.$$

In § 6, we saw that if in $\beta=\pi_1^k\bar{\beta}$ $k \not\equiv 0$, mod 3, then π_1 is the cube of a prime ideal in $K(\beta^{1/3}, \Delta_2^{1/2})$. Now p_5 becomes $\mathfrak{p}_1^3$ in $K(\alpha)$. Then by Dedekind's theorem, we see that p_5 is contained an even number of times in Δ_3, and (for $p_5>3$) in Δ_2. Thus in $K(\Delta^{1/2})$, $p_5=\pi_1$, or $p_5=\pi_1 \cdot \pi_2$. So, in $K(\beta^{1/3}, \Delta_2^{1/2})$, $p_5=\bar{\pi}_1^3$ or $\bar{\pi}_1^3\bar{\pi}_2^3$. That is, when $k \not\equiv 0$, mod 3, we have the case p_5. This is the only way in which p_5 occurs; and no other prime occurs in this way. So for all other primes, $k \equiv 0$, mod 3, and $\bar{\beta}$ has the same cubic character as β. Therefore,

If $k \equiv 0$, mod 3, $\{\beta/\pi\}=\{\bar{\beta}/\pi\}=\omega^i$, $(i=0, 1, 2)$.
If $k \not\equiv 0$, mod 3, $\{\beta/\pi\}=0$.

Finally, then, we have

$$\zeta_K(s)=\prod_p 1/(1-1/p^s) \cdot \prod_\pi 1/[1-\Psi(\pi)/N(\pi)],$$

the first product extending over all rational primes, p, and the second product over all the prime ideals in $K(\Delta_3^{1/2})$.

---

* *American Journal of Mathematics*, Vol. 44, pp. 202 ff.

Differential Equations of Infinite Order with Constant Coefficients.*

By Harold T. Davis.

1. *Introduction.* In a fundamental paper F. Schürer † has discussed the solution of the differential equation of infinite order

$$a_0 u(x) + a_1 u'(x) + a_2 u''(x) + \cdots = f(x), \tag{1}$$

in which the coefficients are assumed to be constants and $f(x)$ is an infinitely differentiable function subject to the condition

$$\lim_{n=\infty} |\, f^{(n)}(x) \,|^{1/n} \leqq L, \qquad (L \text{ finite}). \tag{2}$$

It is clear that equation (1) is of rather general application since its theory is closely associated with the theory of functional equations of the following types:

$$\text{(a)} \quad u(x) + \int_x^\infty \sum_{i=1}^n e^{s_i(x-t)} u(t)\,dt = f(x), \qquad s_i > 0,$$

$$\text{(b)} \quad u(x) + \int_a^b K(t) u(x + ct)\,dt = f(x),$$

$$\text{(c)} \quad u(x + b) + \lambda u(x) = f(x).$$

This relationship is formally exhibited by expanding $u(T)$ in a Taylors series about x,

$$u(T) = u(x) + (T - x)u'(x) + (T - x)^2 u''(x)/2! + \cdots .$$

If we replace T by t and substitute in (a) we get a differential equation in which the coefficients are,

$$a_0 = 1 + \sum_{m=1}^n 1/s_m, \qquad a_i = \sum_{m=1}^n 1/s_m^{i+1}, \qquad i > 0.$$

Similarly, if we set $T = x + ct$, equation (b) is seen to reduce to a differential equation of type (1) with the coefficients

[illegible]

---

* [illegible] 27, [illegible] Colorado.

Letting $T = x + b$, equation (c) reduces to a differential equation with coefficients

$$a_0 = 1 + \lambda, \qquad a_i = b^i/i\,!, \qquad i > 0.$$

One object of the present paper is to discuss the operational solution of equation (1) when the condition (2) imposed by Schürer fails to hold. A second object is to extend theorems of the Heaviside operational calculus to infinite operators by methods which have some claim to novelty since they exhibit a fundamental relationship connecting the Heaviside expansion theorem with the expansion of the operators in Laurent series.*

2. *Expansion of the Resolvant Generatrix in a Laurent Series.* By the generatrix of equation (1) we shall mean the function

$$F(z) = a_0 + a_1 z + a_2 z^2 + \cdots .$$

It is well known that the solution of (1) is given symbolically by the *resolvant generatrix*, $G(z) = 1/F(z)$, in the form

$$u(x) = G(z) \rightarrow f(x),$$

where we employ the symbol $G(z) \rightarrow f(x)$ to signify the function obtained by operating upon $f(x)$ by $G(z)$.†

We first prove the theorem:

THEOREM 1. *If $G_1(z)$ designates a Laurent expansion of $G(z)$ in an annulus formed by two concentric circles about the origin and if $G(z)$ is any other expansion about the origin, then the function*

---

* Since the completion of this paper three studies relating to the same subject have appeared. In a memoir in the *Annales de l'École Normale Supérieure*, Ser. 3, Vol. 46 (1929), pp. 25-53, G. Valiron has investigated the nature of the solution of the homogeneous equation by means of methods originally due to J. F. Ritt. The case of linear systems of differential equations of infinite order with constant coefficients has been the subject of two memoirs by I. M. Sheffer. See the *Annals of Mathematics*, Ser. 2, Vol. 30 (1929), pp. 250-264; *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 281-289. In these latter the degree (Stufe), (see section 3), is limited to the finite case and the approach, differing from that of the present paper, is made through consideration of an equivalent system of linear algebraic equations.

† This follows rigorously from Bourlet's generatrix equation,

$$[G \cdot F] = F \cdot G + \frac{\partial F}{\partial x}\,\frac{\partial G}{\partial z} + \frac{1}{2!}\,\frac{\partial^2 F}{\partial x^2}\,\frac{\partial^2 G}{\partial z^2} + \cdots = 1.$$

Since $F(z)$ is independent of x, $F \cdot G = 1$. See C. Bourlet, "Sur les opérations en général et les équations différentielles linéaires d'ordre infini," *Annales de l'école normale supérieure*, Ser. 3, Vol. 14 (1897), pp. 133-190.

$$U(x) = \{G_1(z) - G(z)\} \rightarrow f(x),$$

where $f(x)$ is arbitrary to within the limits of the existence of the right hand member, is a solution of the homogeneous equation

$$F(z) \rightarrow u(x) = 0.$$

In order to prove this theorem let us assume first that $F(z)$ is of the form

$$F(z) = (z - a)/\phi(z),$$

where $\phi(z)$ has no singularity within or upon the circle of radius $r = a$.

The inverse operator, $G(z) = \phi(z)/(z - a)$, then has the two expansions,

$$G(z) = -\{1 + (z/a) + (z/a)^2 + \cdots\}\,\phi(z)/a,$$
$$G_1(z) = \{1/z + a/z^2 + a^2/z^3 + \cdots\}\phi(z),$$

the first expansion being valid within the circle of radius a and the second in the region exterior to it.

We make the usual interpretation that $1/z \rightarrow f(x) = \int_0^x f(t)\,dt$. Replacing $f(t)$ by its Taylor's expansion we obtain

$$1/z \rightarrow f(x) = (1 - e^{-xz})/z \rightarrow f(x),$$

which expresses the integral as a differential operator.

Similarly it may be proved that

$$1/z^2 \rightarrow f(x) = \{1 - (1 + xz)e^{-xz}\}/z^2 \rightarrow f(x),$$

and, in general,

$$\begin{aligned} 1/z^n \rightarrow f(x) = \{1 - [1 + xz + x^2z^2/2! \\ + \cdots + x^{n-1}z^{n-1}/(n-1)!]e^{-xz}\}/z^n \rightarrow f(x), \\ = \{x^n/n - x^{n+1}z/(n+1) \\ + x^{n+2}z^2/2!(n+2) - \cdots\}/(n-1)! \rightarrow f(x). \end{aligned}$$

Letting $\phi(z) = 1$, and replacing $1/z^n$ in the expansion of $G_1(z) \rightarrow f(x)$ by the expression given above we get:

$$\begin{aligned} G_1(z) \rightarrow f(x) = \{(1 - e^{-xz})/z + a[1 - (1 + xz)e^{-xz}]/z^2 + a^2[1 - (1 + xz \\ + x^2z^2/2!)e^{-xz}]/z^3 + \cdots\} \rightarrow f(x). \end{aligned}$$

It will be clear from the explicit form of $1/z^n \rightarrow f(x)$ that $G_1(0) = \sum_1^\infty a^n x^n/n!\,a = (e^{ax} - 1)/a$, $G_1'(0) = -\sum_1^\infty a^{n-1}x^{n+1}/(n-1)!\,(n+1)a^2$

in general, $G_1^{(n)}(0) = (-1)^n\{e^{ax}[a^n x^n - na^{n-1}x^{n-1} + n(n-1)a^{n-2}x^{n-2} - \cdots \pm n!] - n!\}/a^{n+1}$.

If we replace these values in the expansion

$$G_1(z) = G_1(0) + G_1'(0)z + G_1''(0)z^2/2! + \cdots,$$

and then collect the coefficient of e^{ax}, we see that the operator reduces to the expression,

$$G_1(z) = e^{ax}e^{-xz}(1/a + z/a^2 + z^2/a^3 + \cdots) \\ - (1/a + z/a^2 + z^2/a^3 + \cdots).$$

Taking account of the fact that $z^n e^{-xz} \rightarrow f(x) = f^{(n)}(0)$, we get: *

$$\begin{aligned} G_1(z) \rightarrow f(x) &= e^{ax}\{f(0)/a + f'(0)/a^2 + f''(0)/a^3 + \cdots\} \\ &\quad - \{f(x)/a + f'(x)/a^2 + f''(x)/a^3 \cdots\}, \\ &= e^{ax}\{f(0)/a + f'(0)/a^2 + f''(0)/a^3 + \cdots\} + G(z) \rightarrow f(x). \end{aligned}$$

Similarly for $\phi(z) = z$, we obtain

$$\begin{aligned} G_1(z) \rightarrow f(x) &= e^{ax}\{f(0) + f'(0)/a + f''(0)/a^2 + \cdots\} \\ &\quad - \{f'(x)/a + f''(x)/a^2 + f'''(x)/a^3 + \cdots\} \\ &= e^{ax}\{f(0) + f'(0)/a + f''(0)/a^2 + \cdots\} + G(z) \rightarrow f(x); \end{aligned}$$

and, more generally, for $\phi(z) = z^n$,

$$\begin{aligned} G_1(z) \rightarrow f(x) &= a^{n-1}e^{ax}\{f(0) + f'(0)/a + f''(0)/a^2 + \cdots\} \\ &\quad + G(z) \rightarrow f(x). \end{aligned}$$

Assuming that $\phi(z)$ can be expressed in the form $\phi_0 + \phi_1 z + \phi_2 z^2 + \cdots$, we then derive by addition the result

$$\begin{aligned} G_1(z) \rightarrow f(x) &= [\phi(a)/a]e^{ax}\{f(0) + f'(0)/a + f''(0)/a^2 + \cdots\} \\ &\quad + G(z) \rightarrow f(x). \end{aligned}$$

Since $U(x) = Ce^{ax}$ is obviously a solution of the equation $F(z) \rightarrow u(x) = 0$, the truth of the theorem is demonstrated for the special case assumed above.

Let us next assume that $F(z)$ is of the form

$$F(z) = (z - a_1)(z - a_2)\cdots(z - a_n)/\phi(z),$$

where $a_1, a_2, \cdots, a_n$ are points within an annulus formed by two concentric

* The equivalence of the two expressions $e^{-xz} \rightarrow \{z^n \rightarrow f(x)\}$ and $z^n e^{-xz} \rightarrow f(x)$ is seen at once from Bourlet's formula (*loc. cit.*) if we let $G = e^{-xz}$ and $F = z^n$. Example: $z^{2m}e^{-xz} \rightarrow \sin x = \{z^{2m} - xz^{2m+1} + x^2 z^{2m+2}/2! + \cdots\} \rightarrow \sin x = (-1)^m\{\sin x \cos x - \sin x \cos x\} = 0$.

circles r and R, (the latter having the larger radius), about the origin and $\phi(z)$ has no singularity within or upon R.

Then the generatrix may be written

$$(3) \qquad G(z) = \phi(z)\{1/P'(a_1)(z-a_1) + 1/P'(a_2)(z-a_2) + \cdots + 1/P'(a_n)(z-a_n)\},$$

where we employ the abbreviation

$$P(z) = (z-a_1)(z-a_2)\cdots(z-a_n).$$

Let us designate the expansion of the resolvant within the circle r by $G(z)$ and the expansion in the region exterior to R by $G_1(z)$. We then have from the result of the case of one pole, the expansion

$$(4) \qquad G_1(z) \to f(x) = \sum_{i=1}^{n} [\phi(a_i)/a_i P'(a_i)] e^{a_i x}\{f(0) + f'(0)/a_i + f''(0)/a_i^2 + \cdots\} + G(z) \to f(x).$$

Hence the difference

$$U(x) = \{G_1(z) - G(z)\} \to f(x),$$

is a solution of the homogeneous equation.

Since we have

$d/dz\{P(z)/\phi(z)\}_{z=a_i} = F'(a_i)$, equation (4) can be written neatly as the contour integral

$$G_1(z) \to f(x) = (e^{-xz}/2\pi i)\int_C \{e^{xt}/(t-z)F(t)\}dt + G(z) \to f(x),$$

where C is a path around the zeros of $F(t)$.

The case of multiple poles is treated by a simple device. If the resolvant is

$$G(z) = \phi(z)/(z-a)^r,$$

we may write it in the form

$$G(z) = (\partial^{r-1}/\partial a^{r-1})\phi(z)/(z-a)(r-1)\,!.$$

Hence we have

$$G_1(z) \to f(x) = (\partial^{r-1}/\partial a^{r-1})[1/(r-1)!]$$
$$\times \{[\phi(a)/a]e^{ax}[f(0) + f'(0)/a + f''(0)/a^2 + \cdots]\}$$
$$+ (-1)^r[\phi(z)/a^r]\{1 + rz/a + r(r+1)z^2/a^2\cdot 2\,!$$

[illegible]

The difference between the left hand member and the second term of the right hand side is again seen to be a solution of the homogeneous equation.

It will be pointed out later that the development given above contains essentially a rationale of the Heaviside operational calculus which has played such an important rôle in the theory of electrical circuits.*

It remains for us to discuss the values of the solution and its derivatives at the point $x = 0$. We obtain the following theorem:

THEOREM 2. *If $G_1(z)$ denotes the Laurent expansion of the function*

$$G(z) = \phi(z)/(z - a_1)(z - a_2)\cdots(z - a_n),$$

in the region exterior to the poles $a_1, a_2, \cdots, a_n$, and if $\phi(z)$ is a polynomial of degree $m < n$, then $u_r(x) = z^r G_1(z) \to f(x)$ vanishes at $x = 0$ for $r = 0, 1, 2, \cdots, n - m - 1$.

Proof: Writing $G(z)$ in the form (3) we have from the results of the last theorem,

$$\begin{aligned} u^{(r)}(x) = z^r G(z) \to f(x) = \sum_{i=1}^{n} [\phi(a_i) a_i^r e^{a_i x}/P'(a_i)] \\ \{f(0)/a_i + f'(0)/a_i^2 + \cdots\} \\ - \sum_{i=1}^{n} \phi(z)/P'(a_i) \to \{f^{(r)}(x)/a_i + f^{(r+1)}(x)/a_i^2 + \cdots\}. \end{aligned}$$

Recalling the algebraic identity

$$I(p) = \sum_{i=1}^{n} a_i^p/P'(a_i) = \begin{cases} 1, & p = n - 1, \\ 0, & 0 \leqq p < n - 1, \end{cases}$$

we see that

$$\begin{aligned} u^{(r)}(0) &= \sum_{i=1}^{n} \{\phi_0 + \phi_1 a_i + \cdots + \phi_m a_i^m\} a_i^r \{f(0)/a_i \\ &\quad + f'(0)/a_i^2 + \cdots\}/P'(a_i) - \sum_{i=1}^{n} \sum_{j=0}^{m} \phi_j \{f^{(r+j)}(0)/a_i \\ &\quad + f^{(r+j+1)}(0)/a_i^2 + \cdots\}/P'(a_i), \\ &= \{\sum_{j=0}^{m} \phi_j I(r + j - 1) f(0) + \sum_{j=0}^{m} \phi_j I(r + j - 2) f'(0) + \cdots\} \\ &\quad - \{\sum_{j=0}^{m} I(-1) \phi_j f^{(r+j)}(0) + \sum_{j=0}^{m} I(-2) \phi_j f^{(r+j+1)}(0) + \cdots\}, \\ &= \sum_{j=0}^{m} \phi_j I(r + j - 1) f(0) + \sum_{j=0}^{m} \phi_j I(r + j - 2) f'(0) + \cdots \\ &\quad + \sum_{j=0}^{m} \phi_j I(j) f^{(r-1)}(0). \end{aligned}$$

* See J. R. Carson, *Bulletin of the American Mathematical Society*, Vol. 31 (1926), pp. 43-68; also *Electric Circuit Theory and Operational Calculus*, McGraw-Hill, (1926).

If $r + j - 1 \leqq n - 2$, then $u^{(r)}(0) = 0$. Since j does not exceed m we have $r \leqq n - m - 1$, which is the statement of the theorem.

Corollary. *If $f(x)$ is a function which vanishes together with its first q derivatives at $x = 0$, then $u^{(r)}(0) = 0$ for $r = n + q - m$.*

3. *Application of Borel Summability.* We next seek conditions under which the function

$$u(x) = 1/F(z) \rightarrow f(x) = G(z) \rightarrow f(x)$$

may exist and represent a solution of (1), where $G(z)$ is the Taylor expansion of the resolvant in a circle about the origin.

If a function $f(x)$ is unlimitedly differentiable we shall mean by its degree (stufe) the value L defined by the limit,

$$\lim_{n=\infty} L_n = \lim_{n=\infty} \mid f^n(x) \mid^{1/n} = L.$$

If L is bounded the following theorems can be readily deduced:

A. If $f(x)$ is of degree L, then $f^{(m)}(x)$ is of degree L.

B. If $f_1(x)$ and $f_2(x)$ are of degrees L_1 and L_2 respectively, where $L_2 > L_1$, then $a_1 f_1(x) + a_2 f_2(x)$ will be at most of degree L_2.

C. If $f(x)$ is of degree L, then $f(ax)$ is of degree $\mid a \mid L$.

D. If $f_1(x)$ is of degree L_1 and $f_2(x)$ is of degree L_2, then $f(x) = f_1(x) f_2(x)$ is at most of degree $L_1 + L_2$.*

A thorough treatment of the case where $f(x)$ is a function whose degree is finite may be found in the work of Schürer. It will be clear, however, that this restriction excludes from consideration a large class of equations of which the equation which defines the ψ function may be cited as typical:

$$u(x+1) - u(x) = 1/x,$$

or in terms of a differential operator

$$(e^z - 1) \rightarrow u = 1/x.$$

The following theorem will be found to extend the class of functions to which these operational methods may be applied:

Theorem 3. *If in equation* (1) *$f(x)$ is of the form*

$$f(x) = g(x) + h(x),$$

where $g(x)$ is a function of finite degree L and where $h(x)$ is of the form

$$h(x) = h_1/x + h_2/x^2 + h_3/x^3 + \cdots,$$

---

* [illegible]

then a solution of (1) *exists of the form*

$$u(x)=\int_0^\infty e^{-xt}Q(t)/F(-t)\,dt+1/F(z)\to g(x),$$

where

(5) $$Q(t)=h_1+h_2t+h_3t^2/2\,!+h_4t^3/3\,!+\cdots,$$

provided, (a) $L<1, L\rho<1$, *where* ρ *is the radius of convergence of* $1/F(z)$; (b) *positive values* k, A, *and* M *exist such that* $|\,Q(t)/F(-t)\,|<Ae^{mt}$, *for* $0<k\leqq t\leqq\infty$; *and* (c), $Q(t)/F(-t)$ *is of limited variation in the interval* $0\leqq t\leqq k$.

It will be clear that the series $\sum_{i=1}^{\infty} b_iL^i$ forms a majorante for the series

$$U(x)=1/F(z)\to g(x)=\{b_0+b_1z+b_2z^2+\cdots\}\to g(x),$$

since by hypothesis L is the degree of $g(x)$ and is smaller than the radius of convergence of the series expansion of $1/F(z)$. Since $U(x)$ is thus uniformly convergent we may form the derivative

(6) $$U^{(n)}(x)=1/F(z)\to z^n\to g(x).$$

From (A) above $z^n\to g(x)$ is of degree L and the series $L^n\sum_{i=1}^{\infty} b_iL^i$ forms a majorante for $U^{(n)}(x)$. It thus appears that $U(x)$ is at most of degree L.

To show that $F(z)\to U(x)$ converges uniformly to $g(x)$ let us write $S_n(x)=a_0U(x)+a_1U'(x)+\cdots+a_nU^{(n)}(x)$ and form the difference $A(x)=|\,g(x)-S_n(x)\,|$. Since $1/F(z)=b_0+b_1z+b_3z^3+\cdots$, we have $a_0b_0=1$, $\sum_{i=0}^{k} a_ib_{r-i}=0$, $r>0$. Hence we have the inequalities, $|\sum_{i=0}^{n} a_ib_{n+p-i}\,|<K\rho^p$ and $g^{(n)}(x)<K'L^n$, where K and K' are suitably chosen constants, and thus obtain

$$A(x)\leqq|\sum_{i=0}^{n} a_ib_{n+1-i}\,|\,|\,g^{(n+1)}(x)\,|+|\sum_{i=0}^{n} a_ib_{n+2-i}\,|\,|\,g^{(n+2)}(x)\,|+\cdots,$$
$$\leqq KK'L^{n+1}\rho/(1-L\rho).$$

By hypothesis (a) we have $L\rho<1$ and $L<1$, and hence it follows that $A(x)$ converges uniformly to zero as $n\to\infty$.

We now consider the equation

$$V(x)=1/F(z)\to h(x).$$

Applying the explicit expansion of the resolvant to $h(x)$ we get

$$\begin{aligned} V(x) &= h_1\{b_0/x - b_1/x^2 + 2!\,b_2/x^3 - 3!\,b_3/x^4 + \cdots\} \\ &+ h_2\{b_0/x^2 - 2!\,\,b_1/x^3 + 3!\,b_2/x^4 - 4!\,b_3/x^5 + \cdots\} \\ &+ h_3\{b_0!/x^3 - 3!b_1/2!\,x^4 + 4!\,b_2/2!\,x^5 - 5!\,b_3/2!\,x^6 + \cdots\} \\ &+ \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \end{aligned}$$

In general this series will be divergent, but it is usually summable by the method of Borel. This makes use of the identity $\int_0^\infty e^{-s}s^n ds = n!$ from which we obtain $V(x)$ in the form

$$V(x)=(1/x)\int_0^\infty [e^{-s}\{h_1 + h_2(s/x) + h_3(s/x)^2/2! + \cdots\}/F(-s/x)]ds.$$

Making the substitution $s = tx$, this becomes

$$V(x) = \int_0^\infty e^{-nt}Q(t)/F(-t)\,dt,$$

where $Q(t)$ is defined by (5).

The convergence of this result is easily established under hypotheses (b) and (c) stated in the theorem for we shall then have

$$|\,V(x)\,| < M/x + A \int_0^\infty e^{(m-x)t}dt = M/x + A/(x-m) \text{ for } x > m.$$

The case where $Q(t)/F(-t)$ has a pole of unit order in the interval $(0, \infty)$ is easily disposed of as follows:

Consider the function

$$G(t) = Q(t)/F(-t) - R/(a-t),$$

where R is the residue of $Q(t)/F(-t)$ at the point $t = a$. It is clear that $G(t)$ is regular at $t = a$ so we may consider the function

$$V(x) = \int_0^\infty e^{-tx}G(t)\,dt + \int_0^\infty e^{-tx}R/(a-t)\,dt. \tag{7}$$

The second integral is divergent, but is seen to correspond formally for the case $m = 1$ to the solution of the equation,

$$(a - d/dx)^m \rightarrow W(x) = R/x,$$

which has for its particular integral the function

$$W(x) = (-1)^m R \int_L^a \{e^{a(c-t)}(x-t)^{m-1}/t(m-1)!\,\}dt.$$

[illegible] unit order of the function $Q(t)/F(-t)$, then the solution of the original equation will be

[illegible]

where the R_i are the residues of $Q(t)/F(-t)$ at the points a_i and

$$G(t) = Q(t)/F(-t) - \sum_{i=1}^{m} R_i/(a_i - t).$$

The case of poles of multiplicity m is treated in similar fashion by writing $G(t)$ in the form, $G(t) = Q(t)/F(-t) - \sum_{m=1}^{m} A_m/(a-t)^m$ where the A_m are the Laurent coefficients in the expansion of $Q(t)/F(-t)$. One then adds to the integral $\int_0^\infty e^{-tx} G(t)\,dt$ m functions obtained from $W(x)$ by letting m assume the values 1, 2, $3 \cdots m$ and replacing R by A_m.

Returning to the original problem, we see that the solution of the equation under the restrictions of the theorem is the sum, $u(x) = U(x) + V(x)$.

4. *Examples of the Theory.* Three examples will serve to illustrate the force of the theorems of the paper:

Example 1. Let us consider the equation

$$u(x+1) - u(x) = 1/x,$$

which may be written as $(e^z - 1) \rightarrow u(x) = 1/x$.

The resolvant may obviously be written as follows,

$$(9) \quad \begin{aligned} u(x) &= 1/(e^z - 1) \rightarrow 1/x, \\ &= \{1/z - 1/2 + B_1 z/2! - B_2 z^3/4! + B_3 z^5/6! - \cdots\} \rightarrow 1/x, \end{aligned}$$

where $B_1 = 1/6$, $B_2 = 1/30$, $B_3 = 1/42, \cdots$ are the Bernoulli numbers.

From the well known relation $B_{p+1}/B_p > A(2p+1)(2p+2)$ where A is a constant,* it is clear that series (9) is divergent for all values of x.

But from (8) we see that the solution can be written in the form

$$(10) \qquad U(x) = c + \log x + \int_0^\infty e^{-tx}\{e^t/(1-e^t) + 1/t\}\,dt,$$

which, when $c = 0$, is a well known form for the function $\psi(x) = \Gamma'(x)/\Gamma(x)$.†

The expression given in (9) is readily seen to be the asymptotic expansion of (10). Making the transformation $t = s/x$ and integrating by parts we at once obtain

$$\begin{aligned} U(x) = c + \log x + (1/x)\{&-1/2 \\ &+ \sum_{p=1}^{n} (-1)^p B_p/2p x^{2p-1} + (1/x^{2n+1}) \int_0^\infty e^{-t} \phi^{(2n+1)}(t/x) \end{aligned}$$

* E. Borel, *Leçons sur les séries divergentes*, Paris (1901), p. 24.

† See N. Nielsen, "Handbuch der Theorie der Gammafunktionen," Leipzig (1906), p. 183.

On Generalized Tchebycheff Inequalities in Mathematical Statistics.

By Clarence DeWitt Smith.*

1. *Introduction.* If P is the probability that a variable will deviate from its expected value by an amount as great as t times the standard deviation of the variable, the Tchebycheff inequality states that P is not greater than $1/t^2$ for every t.† This inequality may be regarded as a criterion which places an upper bound on the probability of a deviation of an item from its expected value without regard to the nature of the distribution from which the item is drawn. Bienaymé contributed to the development of such a criterion in a paper ‡ published in 1853. Several contributions have been made in recent years to the generalization of the inequality with a view to obtaining closer inequalities both in the general case where no restrictions are placed on the set of positive values and in cases where certain restrictions are placed on the distribution of the given values. The first important step in generalizing the inequality appeared in a paper § by Karl Pearson in 1919. He obtains the inequality

$$P_{t\sigma} > 1 - M_{2n}/t^{2n}M_2{}^n$$

where M_{2_n} is the $2n^{th}$ moment about the mean and $P_{t\sigma}$ is the probability that a deviation from the mean of the distribution is less than t times the standard deviation if t is a positive number.

The inequality is further generalized in a paper ¶ by B. H. Camp in 1922. By placing a mild restriction on the nature of the frequency function he succeeded in reducing the larger member of the Pearson inequality by about fifty per cent. The main restriction placed on $f(x)$ is that it be a monotonic

* Acknowledgments are due to Professor H. L. Rietz for his helpful suggestions in regard to certain important points discussed in the paper.

† Tchebycheff, "Des Valeurs Moyennes," *Journal de Mathématiques* (2), Vol. 12 (1867), pp. 177-84.

‡ M. Bienaymé, "Considerations à l'appue de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés," *Comptes Rendus*, Vol. 37 (1853), pp. 309-24.

§ Karl Pearson, "On Generalized Tchebycheff Theorems in the Mathematical Theory of Statistics," *Biometrika*, Vol. 12 (1918-19), pp. 284-96.

¶ B. H. Camp, "A New Generalization of Tchebycheff's Statistical Inequality," *Bulletin of the American Mathematical Society*, Vol. 28 (1922), pp. 427-32.

decreasing function of $|x|$ when $|x| \geqq c\sigma$, $c \geqq 0$. With the origin chosen so that zero is at the mean of the distribution he obtains the generalized inequality

$$P_{t\sigma} \leqq \frac{\beta_{2n-2}}{t^{2n}(1+1/2n)^{2n}} \cdot \frac{1}{1+\phi} + \theta.$$

In this inequality $P_{t\sigma}$ is the probability that $|x|$ is as great as $t\sigma$ when t is a positive number,

$$\theta = \frac{\phi \cdot P_{c\sigma}}{1+\phi} \quad \beta_{2n-2} = \frac{M_{2n}}{\sigma^{2n}} \text{ and } \phi = \frac{[2cn/t(2n+1)]^{2n}}{(2n+1)(t/c-1)}.$$

When P_d is the probability that a deviation from the most probable value is as great as $d = tM_n{}^{1/n}$ the above inequality for $c = 0$ is

$$P_d \leqq 1/t^n(1+1/n)^n.$$

This is the inequality due to M. B. Meidell.* Results analogous to those of Pearson and Meidell were published in 1923 in a paper † by Narumi. He represents by $f(x)$ any law of frequency of infinite range and assumes a set of positive deviations from an arbitrary origin by setting

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} [f(x) + f(-x)]\,dx = 1.$$

He also arrives at certain closer inequalities for an increasing function.

It is the main purpose of the present paper to give a further development of the theory and properties of what may be appropriately called generalized Tchebycheff inequalities. We shall give in Section 2 a simple development of the generalized inequality with no restrictions on the nature of the distribution of a set of positive values. The inequality regarded as a proposition of geometry will be discussed in Section 3. In Section 4 we shall derive some closer inequalities, give a discussion of their properties, and a discussion of the effects on the inequality of certain restrictions on the nature of the distribution function. Special functions for which the inequality may be made as close as you please will be discussed in Section 5. The fact that the closeness of the inequality may be improved in certain cases by moving the origin

---

* M. B. Meidell, "Sur un problème du calcul des probabilités et les statistiques mathématiques," *Comptes Rendus*, Vol. 175 (1922), p. 806; also, "Sur la probabilité des erreurs," *Comptes Rendus*, Vol. 176 (1923), p. 280; see also *Skandinavisk Aktuarietidskrift*, årgång V (1922), p. 210.

† Seimatsu Narumi, "On Further Inequalities with Possible Application to Problems in the Theory of Probability," *Biometrika*, Vol. 15 (1923), p. 245. See also M. Alf Guldberg, "Sur le théorème de M. Tchebycheff," *Comptes Rendus*, Vol. 175 (1922), p. 418.

where $\phi(v) = e^v/(1-e^v) + 1/v$.

Consider the remainder

$$R_n = |\ (1/x^{2n+1}) \int_0^\infty e^{-t}\phi^{(2n+1)}(t/x)\,dt\ |.$$

Since we have

$$\phi^{(n)}(v) = \sum_{r=1}^{\infty} A_r e^{rv}/(1-e^{rv})^r + (-1)^n(n-1)!/v^n,$$

where A_r are functions of n alone it is clear that for large values of v, $\phi^{(n)}(v) \sim \sum_{r=1}^{n} (-1)^r A_r$. From this it follows that positive values A and m exist such that we have $|\ \phi^{(2n+1)}(v)\ | < Ae^{mv}$. Hence for $x > m$, we get

$$R_n < (1/x^{2n+1}) \int_0^\infty e^{-t}A\, e^{mt/x}dt = A/x^{2n}(x-m).$$

Consequently it follows that

$$\lim_{|x|=\infty} x^{2n}R_n < \lim_{|x|=\infty} A/(x-m) = 0,$$

which satisfies Poincaré's criterion for the asymptotic expansion of $U(x)$.*

Example 2. The fundamental equation satisfied by the polynomials representing the sums of powers of the natural numbers, $S_n(p) = 1^n + 2^n + \cdots + p^n$, is obviously †

$$S_n(p) - S_n(p-1) = p^n,$$

which may be written symbolically as

$$(1-e^{-z}) \to S_n(p) = p^n.$$

The solution follows immediately:

$$\begin{aligned} S_n(p) &= 1/(1-e^{-z}) \to p^n, \\ &= \{1/z + 1/2 + B_1 z/2! - B_2 z^3/4! + B_3 z^5/6! - \cdots\} \to p^n, \\ &= p^{n+1}/(n+1) + p^n/2 + nB_1 p^{n-1}/2! - n(n-1)(n-2)B_2 p^{n-3}/4! \\ &\quad + n(n-1)(n-2)(n-3)(n-4)(n-5)B_3 p^{n-5}/6! - \cdots, \end{aligned}$$

which is immediately recognized as Bernoulli's famous formula.

Example 3. The Heaviside expansion theorem so useful in the study of electrical circuits is an immediate consequence of equation (4).‡

The problem studied by Heaviside was that of determining the current in an electrical net work when an electromotive force $E(t)$ is impressed on the net work at time $t = 0$. The special feature of this problem is that the unknown function (or functions in the case of multiple net works) denoting

---

* *Acta Mathematica*, Vol. 8 (1886), pp. 295-344.

† See Nielsen, *Traité des nombres de Bernoulli*, Paris (1923), pp. 296-297.

‡ For details and the extensive literature on this theorem see Carson, *loc. cit.*

the current in the net work must vanish together with its derivatives to as high an order as possible at $t=0$. From theorem 2 it is clear that this property is possessed by the solution (4) which, consequently, should be identifiable with the Heaviside expansion theorem.

That this is the case is immediately seen if we change variables from x to t and let $f(t)=E(t)=1$. We then get

$$u(t)=G(z)\to E(t)=\sum_{k=1}^{n} e^{a_k t}/a_k F'(a_k)+1/F(0),$$

which is the Heaviside theorem.

Carson's extension of the Heaviside theorem to an alternating E. M. F. is at once obtained if we let $E(t)=e^{int}$. We thus have

$$u(t)=G(z)\to E(t)=\sum_{k=1}^{n} e^{a_k t}/(a_k-in)F'(a_k)+e^{int}\,1/F(in).$$

The form of solution of the general problem found most useful by Carson is derived in a simple manner from equation (4).

Making use of the fact that in the development

$$1/F(z)=b_0+b_1z+b_2z^2+\cdots,$$

the coefficients may be calculated from the formulas

$$b_0=1/F(0),\qquad b_m=-\sum_{k=1}^{n} 1/a_k^{m+1}F'(a_k),\qquad m>0,$$

where the a_k are the roots of $F(z)=0$, we may write

$$\begin{aligned}
u(t)=&\{\sum_{k=1}^{n} e^{a_k t}E(0)/a_kF'(a_k)+\sum_{k=1}^{n} e^{a_k t}E'(0)/a_k^2F'(a_k)\\
&+\sum_{i=1}^{n} e^{a_k t}E''(0)/a_k^3F'(a_k)+\cdots\}+E(t)/F(0)\\
&-\sum_{k=1}^{n} E'(t)/a_k^2F'(a_k)-\sum_{k=1}^{n} E''(t)/a_k^3F'(a_k)-\cdots,\\
=&\sum_{j=1}^{\infty}\sum_{k=1}^{m}\{e^{a_k t}E^{(j-1)}(0)/a_k^jF'(a_k)-E^{(j-1)}(t)/a_k^jF'(a_k)\}\\
&+E(t)/F(0)+\sum_{k=1}^{n} E(t)/a_kF'(a_k),\\
=&\int_0^t\{\sum_{k=1}^{n} e^{a_k(t-s)}E(s)/F'(a_k)\}ds+E(t)/F(0)+\sum_{k=1}^{n} E(t)/a_kF'(a_k).
\end{aligned}$$

This is easily seen to be equivalent to the expansion

$$u(t)=z\to\int_0^t\{\sum_{k=1}^{n} e^{a_k(t-s)}/a_kF'(a_k)+1/F(0)\}E(s)\,ds,$$

which is Carson's formula.

WATERMAN INSTITUTE,
INDIANA UNIVERSITY.

will be discussed in Section 6. In Section 7 we shall discuss certain types of distributions for which the difference between the members of the inequality attains a minimum.

2. *Development of the Inequality from Markoff's Lemma.* The inequality has been developed by different authors from somewhat different viewpoints. Tchebycheff assumed a set of deviations of variables from their expected values. Pearson and Camp assumed a set of deviations from the arithmetic mean and considered moments of even order. Meidell assumed a set of deviations from the most probable value. Narumi assumed a set of positive deviations from an arbitrary origin. Hence it should be of interest at this point to develop a generalized inequality which gives rise to several important known forms of the generalized Tchebycheff inequality in a very simple way. We shall begin by giving a proof of a generalized form of the Markoff Lemma * which may be stated in the following modified form.

If a variable U takes any set of positive values, $u_1, u_2, \cdots, u_k$ with the corresponding probabilities, $p_1, p_2, \cdots, p_k$ respectively, and if M_n is the expected value of U^n, then the probability P_d that U is greater than d is not greater than M_n/d^n where d is a positive number.

To prove this generalized form of the lemma let $u_c < d \leqq u_{c+1}$. Then the probability that U is greater than d is $P_d = p_{c+1} + p_{c+2} + \cdots + p_k$.

$$M_n = U_1{}^n p_1 + U_2{}^n p_2 + \cdots + U_c{}^n p_c + U^n{}_{c+1} p_{c+1} + \cdots + u_k{}^n p_k$$
$$\geqq d^n(p_{c+1} + p_{c+2} + \cdots + p_k) = d^n P_d.$$

Hence

(1) $$P_d \leqq M_n/d^n.$$

Now let us consider the forms of the inequality which result from certain assigned values of n and d. First let $n = 2$ and $d = tM_2{}^{1/2}$ where M_2 is the second moment of deviations of a variate from its expected value. Then,

(2) $$P_d \leqq 1/t^2,$$

which is the Tchebycheff inequality for a single variate. Next let n be an even number and $d = tM_2{}^{1/2}$ where M_2 is the second moment of deviations of a variate from the arithmetic mean. Then $M_2{}^{1/2}$ is the standard deviation σ, and

(3) $$P_d \leqq M_n/(t^n\sigma^n)$$

gives a bound analogous to that of Pearson. Finally, if $d = tM_n{}^{1/n}$,

(4) $$P_d \leqq 1/t^n.$$

This inequality is analogous to that of Narumi. Hence we have shown how the generalized inequality (1) gives rise to several of the important known inequalities in a rather simple way.

3. *The inequality regarded as a proposition of geometry.* The theorem of Tchebycheff and its generalizations may be regarded as criteria which hold between the area and the moments of area under a curve. If $y = f(x)$ is a single valued continuous function of x, then for a certain area in the first quadrant the theorem may be stated as an inequality which holds between the moments of area under the curve and the area under a given part of the curve. When so stated it may be put in the following form.

If we take a unit of area in the first quadrant bounded by a curve $y = f(x)$, the x-axis, and the lines $x = a$ and $x = k$, $(k > a)$, then the nth moment M_n, $(n > 0)$, of this unit area about the y-axis is not less than the product of d^n and the part of this area to the right of the line

$$x = d, \ (M_n{}^{1/n} < d \leqq k).$$

4. *On some further relatively close inequalities.* It is shown in Section 7 for any distribution of infinite range that the two members of the Tchebycheff inequality will approach arbitrarily near to equality as d is increased indefinitely. But the inequality is not in general close to an equality when d is a finite number of the order of $M_n{}^{1/n}$ except for specially designed functions. It would seem to be of interest to obtain a closer inequality connecting d, P_d, and the nth power of the values of the variable applicable to a wide class of functions, and if possible to obtain an inequality that will in general approach an equality at the ends of a bounded distribution.

Consider again the increasing sequence of positive numbers $u_1, u_2, u_3, \cdots, u_k$ which are the values taken by a variable U with the respective probabilities $p_1, p_2, p_3, \cdots, p_k$. Following the notation in Section 2, let d be chosen so that $u_c < d \leqq u_{c+1}$ and let $M_n = \sum_{i=1}^{k} u_i{}^n p_i$ be the expected value of U^n. Then it follows that

$$\sum_{i=c+1}^{k} u_i{}^n p_i \geqq d^n \sum_{i=c+1}^{k} p_i, \qquad (n > 0).$$

But $\sum_{i=c+1}^{k} p_i = P_d$, the probability that a value of U taken at random is greater than d. Hence,

$$P_d \leqq (1/d^n) \sum_{i=c+1}^{k} u_i{}^n p_i. \tag{5}$$

If $f(x) \geqq 0$ is a probability function that is continuous and integrable from d to k, this inequality takes the form

$$P_d \leqq (1/d^n) \int_d^k x^n f(x)\, dx, \tag{6}$$

where

$$P_d = \int_d^k f(x)\, dx.$$

Next let $x_m = \sum_{i=c+1}^{k} u_i p_i / \sum_{i=c+1}^{k} p_i$ be the mean value of $u_{c+1}, u_{c+2}, \cdots, u_k$, weighted with their respective probabilities. Since it is well known[*] that for $n \geqq 1$,

$$\sum_{i=c+1}^{k} u_i^n p_i / \sum_{i=c+1}^{k} p_i \geqq \left(\sum_{i=c+1}^{k} u_i p_i / \sum_{i=c+1}^{k} p_i\right)^n = x_m^n,$$

we have

$$P_d \leqq (1/x_m^n) \sum_{i=c+1}^{k} u_i^n p_i. \tag{7}$$

Similarly for the continuous probability function $f(x)$ we have

$$P_d \leqq (1/x_m^n) \int_d^k x^n f(x)\, dx. \tag{8}$$

It seems desirable to emphasize that the inequality (7) is in general a closer inequality than (5) because $x_m > d$ for the increasing sequence. When $f(x)$ is an increasing function throughout the interval d to k, x_m is greater than $\tfrac{1}{2}(k + d)$. Therefore for a function $f(x)$ that increases throughout the interval d to k it follows that

$$P_d \leqq [2/(d + k)]^n \int_d^k x^n f(x)\, dx. \tag{9}$$

In the inequalities (7), (8), and (9) it should be noted that $n \geqq 1$. In all cases where the function $f(x)$ increases from d to k, (9) holds and is a closer inequality than (6). To compare the relative values of the actual probability and the values given by (6) and (9) we may take as a simple case the function $f(x) = x^2$. Let $2 \int_0^k x^2 dx = 1$, $n = 4$, and we have

d	P	(6)	(9)
1.1	.112	.1223	.1135
1.5	0	0	0

A comparison of these values shows the very close approximation to P as given by (6) and (9). In applying the inequality (9) it is well to note that there is a limitation on the value of x in terms of n and M_n in the case

of a non-decreasing frequency function. For if we assume a unit area bounded by $y = f(x)$, the x-axis, the y-axis, and the line $x = k$ where $f(x)$ is a non-decreasing function in the interval $x = 0$ to $x = k$ then *

$$k \leqq (1 + n)^{1/n} M_n^{1/n}; \qquad n \geqq 1. \tag{10}$$

To establish this inequality let

$$\bar{y} = (1/k) \int_0^k y dx = 1/k$$

be the mean ordinate of the curve $y = f(x)$. For the nth moment of the rectangle $OBCR$, where $OB = 1/k$, we have

$$M'_n = (1/k) \int_0^k x^n dx = k^n/(n + 1).$$

Hence for the constant distribution we have

$$k = (1 + n)^{1/n} (M'_n)^{1/n}.$$

Since $y = f(x)$ is a non-decreasing function in the interval $0 \leqq x \leqq k$ it is geometrically fairly obvious for $n \geqq 1$, that the nth moment of the area under the curve $y = f(x)$ is not less than the nth moment of the rectangle. That is, $M_n^{1/n} \geqq (M'_n)^{1/n} = k(1 + n)^{-1/n}$, or $k \leqq (1 + n)^{1/n} M_n^{1/n}$. It now follows that d for non-decreasing functions is restricted in accord with the inequality, $M_n^{1/n} < d \leqq k \leqq (1 + n)^{1/n} M_n^{1/n}$.

While the right member of the inequality (5) sets an upper bound for the probability P_d that seems ordinarily more difficult to compute than $P_d = \sum_{i=c+1}^{k} p_i$ itself, nevertheless (5) serves as a starting point for the development of certain useful inequalities. By sacrificing something as regards the smallness of the upper bound in the inequality (5) we may write the inequality in the form

$$P_d \leqq (1/d^n) [\sum_{i=1}^{k} u_i^n p_i - \psi_d], \tag{11}$$

where

$$\psi_d \leqq \sum_{i=1}^{c} u_i^n p_i (u_c < d \leqq u_{c+1}).$$

For a simple case, we may take

$$\psi_d = u_1^n \sum_{i=1}^{c} p_i = u_1^n (1 - P_d),$$

since $u_1 < u_2 < \cdots < u_c$.

* Narumi, *loc. cit.*, p. 249.

Substitution of this value in (11) gives

$$(12) \qquad P_d \leqq \Big(\sum_{i=1}^{k} u_i{}^n p_i - u_1{}^n\Big)/(d^n - u_1{}^n).$$

Similarly, for the case of a continuous function $y = f(x)$ with a range a to k $(a < k)$, we write in place of (11),

$$(13) \qquad P_d \leqq (1/d^n)\Big[\int_a^k x^n f(x)\,dx - \psi_d\Big],$$

where

$$\psi_d \leqq \int_a^d x^n f(x)\,dx.$$

By taking $\psi_d = a^n \int_a^d f(x)\,dx = a^n(1 - P_d)$ we have

$$(14) \qquad P_d \leqq (M_n - a_n)/(d^n - a^n).$$

In other cases we may take $\psi_d = \int_a^d x^n y\,dx$, where $y = f_1(x)$ is a function of convenient form and such that

$$\int_a^d x^n f_1(x)\,dx \leqq \int_a^d x^n f(x)\,dx,$$

For example, it may lead to useful results in certain cases to take $y = f_1(x)$ as a suitably selected tangent or chord of the curve $y = f(x)$, the selection to be made depending on the nature of the function $f(x)$ in the interval $x = a$ to $x = k$.

Although such a direct treatment of the problem is possible we have found it convenient to proceed with a less direct plan for obtaining a closer inequality than $P_d \leqq (M_n/d^n)$ by first changing somewhat the form of the inequality (6). For the present we shall limit our considerations to the continuous function $f(x) = 0$ in the interval $0 \leqq x \leqq k$ such that $\int_0^k f(x)\,dx = 1$. While we take x positive, it is fairly obvious that any frequency curve $y = g(x)$ in which x takes negative as well as positive values may be included by taking $f(x) = g(x) + g\,(-x)$. By using moments $M_{2r} = \int_0^k x^{2r} f(x)\,dx$ of even order $2r$ we obtain from (6),

$$(15) \qquad P_d \leqq \Big[M_{2r} - \int_0^d x^{2r} f(x)\,dx\Big]/d^{2r} = 1/t^{2r} - R,$$

where $r \geqq 1$, $d = t(M_{2r})^{1/2r}$, and $R = (1/d^{2r})\int_0^d x^{2r} f(x)\,dx$. Now we may

[illegible]

Let $\phi(x)$ be such that $(d/dx)[\phi(x)] = f(x)$ and let $P_x = \int_x^k f(x)\,dx$. Then $\phi(x) = -P_x$, $\int_0^d x^{2r} f(x)\,dx = -d^{2r}P_d + 2r\int_0^d x^{2r-1}P_x dx$, and

$$R = -P_d + (2r/d^{2r})\int_0^d x^{2r-1}P_x dx. \tag{16}$$

From (15) and (16),

$$2r\int_0^d x^{2r-1}P_x dx \leqq M_{2r}. \tag{17}$$

Since $(d/dx)P_x = -f(x)$ is negative the second derivative of P_x is positive when $f(x)$ is a decreasing function. Let us consider the case where $f(x)$ is a decreasing function in the interval $0 \leqq x \leqq d$. In this case the curve $y = P_x$ is concave upward. The equation of a tangent to the curve $y = P_x$ at $x = \theta d$, $(0 \leqq \theta \leqq 1)$ is

$$y = -xf(\theta d) + P_{\theta d} + \theta d f(\theta d). \tag{18}$$

Since $y = P_x$ is concave upward it does not cross this tangent and the inequality (17) remains valid if we replace P_x by the right member of (18). This gives after integration,

$$-[2r/(2r+1)]\,d^{2r+1}f(\theta d) + d^{2r}[P_{\theta d} + \theta d f(\theta d)] \leqq M_{2r},$$

or

$$P_{\theta d} + df(\theta d)[\theta - 2r/(2r+1)] \leqq M_{2r}/d^{2r}. \tag{19}$$

The closeness of this inequality depends on the selection of θ. To find a value of θ which would make the non-negative difference between the members of (19) a minimum we write

$$V = M_{2r}/d^{2r} - P_{\theta d} - df(\theta d)[\theta - 2r/(2r+1)],$$

and seek the value of θ which makes

$$dV/d\theta = df(\theta d) - df(\theta d) - d^2 f'(\theta d)\,[\theta - 2r/(2r+1)] = 0.$$

This equation is satisfied by $\theta = 2r/(2r+1)$. The second derivative of V is positive for this value of θ and the condition for a minimum is satisfied. If we let $\theta = 2r/(2r+1)$ in (19),

$$P_{2rd/\,2r+1} \leqq M_{2r}/d^{2r}.$$

A simple transformation gives

$$P_d \leqq M_{2r}/d^{2r}(1 + 1/2r)^{2r}.$$

If $d = t(M_{2r})^{1/2r}$,

$$P_d \leqq 1/t^{2r}(1 + 1/2r)^{2r}. \tag{20}$$

This is the inequality of Meidell and of B. H. Camp for his $c = 0$. We have proved that this inequality is closer to an equality than any other derivable by using in place of P_x in (17) the function represented by a tangent to the curve $y = P_x$ at any point of the interval $x = 0$ to $x = d$.

Next let us consider the case in which $f(x)$ is an increasing function. In this case we know from (10) that the range 0 to k for consideration is limited to $t \leqq (1 + 2r)^{1/2r}$. Now the second derivative of P_x is negative and the curve $y = P_x$ is concave downward. Consider the function represented by a chord which connects the points on $y = P_x$ for which $x = 0$ and $x = d$. The equation of this chord is

$$y = (x/d)(P_d - 1) + 1. \tag{21}$$

When we substitute the right member of (21) for P_x in (17),

$$(2r/d)(P_d - 1)\int_0^d x^{2r}dx + 2r\int_0^d x^{2r-1}dx \leqq M_{2r}.$$

After integration this equation reduces to

$$P_d \leqq (M_{2r}/d^{2r})(1 + 1/2r) - 1/2r.$$

When $d = t(M_{2r})^{1/2r}$,

$$P_d \leqq (1/t^{2r})(1 + 1/2r) - 1/2r. \tag{22}$$

On comparing the upper bound in (22) with the upper bound set by the inequality of Narumi * for an increasing function, which is

$$[1 - b^n/(n+1)]/[t^n - b^n/(n+1)]; \quad t \geqq b,$$

it seems that they are equal when $t = b$ and that the upper bound in (22) is smaller when $t > b$.

Let us consider next the case in which $y = f(x)$ is an increasing function when $0 < x < h$ $(0 < h < k)$, and a decreasing function when $x \geqq h$. Then the function $Y = P_x$ will be represented by a curve which is concave downward from $x = 0$ to $x = h$ and concave upward beyond $x = h$. We may substitute for $\int_0^d x^{2r-1}P_x dx$ in (17) the moment under the horizontal line AB from zero to h plus the moment under the tangent at $x = \theta d$ from h to d where as above $0 \leqq \theta \leqq 1$. We will then have from (17)

$$2r\int_0^h [illegible]$$

$$P_h h^{2r} + \{-[2rd^{2r+1}/(2r+1)]f(\theta d) + d^{2r}[P_{\theta d} + \theta d f(\theta d)]$$

$$+ [2rh^{2r+1}/(2r+1)]f(\theta d) - h^{2r}[P_{\theta d} + \theta d f(\theta d)]\} \leqq M_{2r},$$

or

$$(23) \quad P_{\theta d} + \theta d f(\theta d) - [2r/(2r+1)]\,[(d^{2r+1} - h^{2r+1})/(d^{2r} - h^{2r})] f(\theta d) \leqq (M_{2r} - h^{2r}P_h)/(d^{2r} - h^{2r}).$$

We may ask for the value of θ which will make the non-negative difference between the two members of (23) a minimum. Let the difference be represented by

$$V = (M_{2r} - h^{2r}P_h)/(d^{2r} - h^{2r}) - P_{\theta d} - \{\theta d - [2r/(2r+1)]\,[(d^{2r+1} - h^{2r+1})/(d^{2r} - h^{2r})]\} f(\theta d).$$

Then if we set $dV/d\theta = 0$, we find that the value of θ which satisfies the condition for a minimum is

$$(24) \quad d' = \theta d = [2r/(2r+1)]\,[(d^{2r+1} - h^{2r+1})/(d^{2r} - h^{2r})].$$

By substituting in (23) the value of θd given by (24),

$$P_{d'} \leqq (M_{2r} - h^{2r}P_h)/[(d'/\theta)^{2r} - h^{2r}].$$

We may let $h = c(M_2)^{1/2}$, and if $y = f(x)$ is the distribution of absolute values from an arithmetic mean, we have $M_2^{1/2} = \sigma$, where σ is the standard deviation. Then if $d' = t\sigma$ we have

$$(25) \quad P_{d'} \leqq (\beta_{2r-2} - c^{2r}P_{c\sigma})/[(t/\theta)^{2r} - c^{2r}],$$

where

$$\beta_{2r-2} = (M_{2r}/\sigma^{2r}).$$

Here θ is determined from the equation

$$t = [2r/(2r+1)]\,[t^{2r+1} - (c\theta)^{2r+1}]/\theta[t^{2r} - (c\theta)^{2r}].$$

Now we know that $P_{c\sigma}$ is greater than $P_{t\sigma}$ and if we substitute $P_{t\sigma}$ for $P_{c\sigma}$ in (25),

$$(26) \quad P_{t\sigma} \leqq \beta_{2r-2}/(t/\theta)^{2r}.$$

If we set $c = r = 1$ and $t = 2$ in (26) we get $P_{2\sigma} \leqq .133$.

Referring again to the function $y = P_x$ with the assumptions which lead to (23) we find that the equation of the chord $A'B$ from $x = 0$ to $x = h$ is

$$y = (1/h)[x(P_h - 1) + h].$$

If we integrate from $x = 0$ to $x = h$ using this function, the first integral in the development which lead to (23) will be

$$2r \int_0^h (1/h)[x(P_h - 1) + h]\, x^{2r-1} dx = h^{2r}(2rP_h + 1)/(2r+1).$$

Proceeding as above we obtain in place of (25),

$$(27) \qquad P_{t\sigma} \leqq \{\beta_{2r-2} - c^{2r}[(2rP_{c\sigma}+1)/(2r+1)]\}/[(t/\theta)^{2r} - c^{2r}].$$

We may again substitute $P_{t\sigma}$ for $P_{c\sigma}$ as in the preceding case and

$$(28) \qquad P_{t\sigma} \leqq [\beta_{2r-2} - c^{2r}/(2r+1)]/[(t/\theta)^{2r} - c^{2r}/(2r+1)],$$

where θ is determined as in (25). If $c = r = 1$ and $t = 2$ in (28) we have $P_{2\sigma} \leqq .092$. The inequality of Camp for the same values gives $P_{2\sigma} \leqq .143$.

5. *On close inequalities for special functions.* Narumi * states that functions exist for which the difference between the two members of Pearson's inequality may become arbitrarily small. In a recent paper † Bortkiewicz presents a discussion of the Markoff Lemma in which he gives a method of determining a distribution for which the difference between the members of the inequality stated in the lemma may be made equal to a previously assigned small number. A similar method may be applied to a generalized Tchebycheff inequality. In the present section we shall present a discussion of a method applicable to the generalized Tchebycheff inequality

$$P_d \leqq M_n/d^n,$$

in such a way as to bring out the point that the equality is not realized except as a limiting case.

Assume a distribution of positive varieties composed of the parts a and $1 - a$. Let the variates which belong to $1 - a$ be distributed in any convenient manner from the origin to a point at a distance d_1 from the origin. Then place the variates which belong to a, at a distance $d > d_1$. If M_n' is the nth moment of that part of the distribution represented by $1 - a$,

$$M_n = M_n' + ad^n, \text{ or } M_n/d^n = (M_n'/d^n) + a.$$

It is now fairly obvious that as d_1 is decreased we can make M_n' approach zero and M_n/d^n will approach a. But $P_{d-\epsilon} = a$ for $\epsilon < d - d_1$ and $M_n/(d-\epsilon)^n$ approaches M_n/d^n as ϵ becomes arbitrarily small. Hence the difference, $M_n/(d-\epsilon)^n - P_{d-\epsilon}$, becomes arbitrarily small for sufficiently small values of ϵ.

The following distributions have been found to give rather close values for the inequality, and they seem to suggest probable forms of a continuous

---

* *Loc. cit.*, p. 245.

† [illegible]

function which would have a ponit at which the inequality may be expected to become fairly close.

p_i	.488	.005	.004	.0025	.0005	0	.0005	.0025	.004	.005	.488
x_i	0	1	2	3	4	5	6	7	8	9	10

Let $tM_2^{1/2} = 10 - \epsilon$. Then $P = .488$ and $1/t^2 \leqq .4965$.

p_i	0	.0005	.0006	.0007	.0008	.0009	.9965
x_i	0	1	2	3	4	5	6

Let $tM_2^{1/2} = 6 - \epsilon$. Then $P = .9965$ and $1/t^2 \leqq .9977$.

In these illustrations ϵ is a positive number small at pleasure.

6. *On certain changes in the Tchebycheff inequality produced by moving the origin.* It is well known that the Tchebycheff inequality and certain of its generalizations are valid for positive deviations from any origin. However the degree of approach toward equality of the two members of the inequality may depend upon the origin from which deviations are measured. If $|x_1|, |x_2|, \cdots |x_k|$, are non-negative deviations from any origin it is well known that

$$P_d \leqq M_n/d^n, \tag{29}$$

where P_d is the probability that an $|x|$ taken at random from the population is as great as d.

In considering the degree to which the right hand member of (29) can be made to approach the left hand member by a suitable selection of an origin from which to measure deviations we shall assume that the original origin is such that we have a set of non-negative deviations. We propose first to consider the question of moving the origin so as to change the value of each deviation by an amount h where h may take positive values and negative values such that $|h|$ is not greater than the smallest deviation from the original origin. Now if P_d is the probability of a deviation as great as d from the original origin, $P_d = P'_{d+h}$ where P'_{d+h} is the probability of a deviation as great as $d + h$ from the new origin. Then in order to find the origin for which the right hand member of the inequality

$$P'_{d+h} \leqq \sum_{i=a}^{k} (|x_i| + h)^n/(d+h)^n$$

shall approach as near as possible to the left hand member we seek the value of h which will make

$$V_1 = \sum (|x_i| + h)^n/(d+h)^n \tag{30}$$

a minimum. Such a value of h will satisfy the equation

$$(31) \qquad dV_1/dh = \frac{n(d+h)[\sum(|x_i|+h)^{n-1}] - n\sum(|x_i|+h)^n}{(d+h)^{n+1}} = 0.$$

If $n = 2$ in (31),

$$(32) \qquad h = (\Sigma\,|x_i|^2 - d\Sigma\,|x_i|)/(Nd - \Sigma\,|x_i|).$$

If the original set of values were the absolute values of deviations from an arithmetic mean we may well write (32) in the form

$$(33) \qquad h = (\sigma^2 - dD)/(d - D),$$

where σ is the standard deviation and D is the mean deviation from the arithmetic mean. An examination of (33) shows that h is ordinarily positive for $d < \sigma^2/D$ and negative for $d > \sigma^2/D$; for, in general, $d - D$ is positive since, in general, $d > \sigma > D$.

To illustrate by a simple numerical case, let us take x as the absolute values of deviations from an arithmetic mean with the following distribution:

Frequency	2,	6,	d,	2,
$\lvert x_i \rvert$	0,	1,	1.3,	2.

Here $\sigma^2 = 1.4$, $d = 1.3$, $D = 1$, and $h = 1/3$. Next consider the distribution

Frequency	0	2	6	d	2,
$\lvert x_i \rvert$	0	1	2	2.3	3.

In this case $\sigma^2 = 4.4$, $d = 2.3$, $D = 2$, and $h = -2/3$, $P_d = .832$, and $P'_{d+h} = .812$. These results show the amount of improvement in the value of the approximation in this very simple case.

We may perhaps form a useful conception of the general nature of certain frequency distributions concerning which the problem of changing the origin is likely to arise by considering a special case for illustration. Assume a frequency curve $y = f(x)$ of considerably less range in the negative direction from some convenient origin than in the positive direction. If d is not less than the range of $y = f(x)$ in the negative direction the probability P_d that a datum taken at random from the distribution will deviate from 0 as much as d in absolute value is the same as the probability P'_{d+h} that the same datum will deviate as much as $d + h$ in absolute value from an origin $0'$ located at a distance h to the left of 0. Thus the problem of changing the origin is likely to arise in a different way from that discussed

function with the origin at some such point as the arithmetic mean or the mode. Then suppose we seek to move the origin a distance h so as to make the right-hand member of the Tchebycheff inequality

$$P'_{d+h} \leqq M_n/(d+h)^n$$

approach as near as possible to the left-hand member.

This problem differs from the one considered above in that we now seek a value of h which will make

$$V = \Sigma(|x+h|)^n/(d+h)^n, \tag{34}$$

a minimum. If n is an even number the required value of h satisfies the equation

$$dV/dh = [n(d+h)\Sigma(x+h)^{n-1} - n\Sigma(x+h)^n]/(d+h)^{n+1} = 0.$$

When $n = 2$ in this equation

$$h = (\Sigma x^2 - d\Sigma x)/(nd - \Sigma x) = (M_2 - dM_1)/(d - M_1). \tag{35}$$

If the first origin is taken at the arithmetic mean, $M_1 = 0$ and (35) becomes

$$h = \mu_2/d. \tag{36}$$

Since h in (36) is positive we know the origin which we seek is to the left of the mean, when the distribution is represented as we have assumed.

To illustrate with a continuous function we take $y = x(10-x)^3$. The centroid of area is at $\bar{x} = 3\tfrac{1}{3}$. The second moment about the centroid is $\mu_2 = 3.93$. When $d = 5$, $h = .786$. The second moment about the new origin is $M_2 = 4.538$, and the deviation is $d + h = 5.786$. We then have $P_d \leqq .16$, and $P'_{d+h} \leqq .13$. We may compare .13 with .16 to note the decrease in the larger member of the inequality due to changing the origin. It is interesting to note that in these illustrations the Tchebycheff inequality gives the best estimate of the value of the probability when the origin is quite near the lower bound of the distribution.

7. *On minimal values of the differences between the members of the inequality.* In the inequality $P \leqq 1/t^n$ it is fairly obvious from very elementary considerations that P and $1/t^n$ each approaches zero as $t \to \infty$. Hence for any distribution of infinite range there exists a deviation from the origin for which the difference between the two members of the inequality is arbitrarily small. But tests with actual distributions indicate that the differences between the two members of the inequality as a rule are not very small except for large values of t. The question may well be raised

as to whether there is a finite value of the variable within the range of its values for which the difference between the two members of the inequality is a minimum. It seems that some distributions are of such a nature that the difference decreases as the deviation d from the origin increases while others have places within the distribution at which the difference is a minimum. In the present section we shall give the results of investigating this question for certain special functions.

To consider this question let us assume a distribution function $f(x)$, continuous and integrable from $-a$ to a, such that $\int_{-a}^{a} f(x)\,dx = 1$. If $x_1 = tM_n^{1/n}$ is a deviation from the mean value of x, we know that

$$P_{x_1} \leqq M_n/x_1^n, \tag{37}$$

where P_{x_1} is the probability that a deviation from the mean is as great as $tM_n^{1/n}$. The inequality (37) may be written

$$\int_{-a}^{x_1} f(x)\,dx + \int_{x_1}^{a} f(x)\,dx \leqq (1/x_1^n) \int_{-a}^{a} x^n f(x)\,dx. \tag{38}$$

We would now ask if there are intervals on x for certain functions $f(x)$ which contain a point at which the difference between the members of (38) is a minimum, and if so what is the nature of the function and where are the intervals.

Let $f(x)$ be positive and symmetrical about the y-axis. Then from (38) we may write the difference which is to be a minimum in the form

$$V = (1/x_1^n) \int_0^a x^n f(x)\,dx - \int_{x_1}^{a} f(x)\,dx. \tag{39}$$

To examine V for a minimum, we have

$$dV/dx_1 = -(n/x_1^{n+1}) \int_0^a x^n f(x)\,dx + f(x_1) = 0,$$

or

$$x_1^{n+1} f(x_1) = n \int_0^a x^n f(x)\,dx. \tag{40}$$

If $f(x)$ is an increasing function, d^2V/dx_1^2 is positive for $n > 0$. Hence a real solution of (40) for x_1 within the interval $M_n^{1/n} \leqq x_1 \leqq a$ corresponds to a minimum for V when $f(x)$ is an increasing function at $x = x_1$. If $f(x)$ is a decreasing function d^2V/dx_1^2 is not necessarily positive and a test for a minimum must be made in each case.

seek the value of x_1 that will make V in (39) take a minimal value. In the following cases the ordinary tests for a minimum show that V takes a minimal value at the points indicated.

(a). First assume $f(x) = k$, a constant.

From $2\int_0^a k dx = 1$, we have $k = 1/2a$. We then find $\sigma = a/3^{1/2}$.

From (40), $kx_1^{n+1} = nk \int_0^a x^n dx$ and $x_1^{n+1} = na^{n+1}/(n+1)$.

Let $n = 2$. Then $x_1 = (a/3)(18)^{1/3} > \sigma$.

A minimal value of V exists at this point and we have a very simple function for which the difference between the two members of the inequality attains a minimum at a point fairly near the origin.

(b). Next assume $f(x) = k(1 - x^2)$, and $a = 1$.

From $2k \int_0^1 (1 - x^2) dx = 1$, $k = 3/4$. We find $\sigma = (.2)^{1/2}$.

There seems to be no point at which the difference between the two members of $P \leqq M_n/x_1^n$ attains a minimum. However, in this case $f(x)$ is a monotonic decreasing function of x and the inequality of B. H. Camp applies. To test this inequality for a minimum we seek a root of

$$x_1^{n+1}(1 + 1/n)^n(1 - x_1^2) = 2n/(n+1)(n+3),$$

which is between $(.2)^{1/2}$ and 1. Let $n = 2$ and

$$x_1^5 - x_1^3 + 16/135 = 0.$$

This equation has a root between .55 and .56 which corresponds to the minimal value of V.

We have thus found a particular function $f(x)$ for which the closer approximation of Camp would attain a minimum at a value of x which is roughly 11/20 of the value of x at the upper bound of the distribution. It is a fact of some interest that we have found simple functions for which the minimal value of the difference between the members of the inequality exists for values of x_1 in the interval $M_n^{1/n} < x_1 < a$ in case the inequality of Camp is used, but that no minimum exists for the difference if the inequality $P \leqq M_n/x_1^n$ is used.

(c). Take $f(x) = kx^{2b}$ where b is a positive integer $\geqq 1$. From $2\int_0^a kx^{2b} dx = 1$, $k = (2b+1)/2a^{2b+1}$. We find $\sigma = a[(2b+1)/(2b+3)]^{1/2}$.

From (40) when $n = 2b$,

$$x_1^{2b+1}(x_1^{2b}) = 2b \int_0^a x^{4b} dx, \text{ and } x_1 = a[2b/(4b+1)]^{1/(4b+1)}.$$

Now $x > \sigma$ if

$$[2b/(4b+1)]^{1/(4b+1)} > [(2b+1)/(2b+3)]^{\frac{1}{2}}.$$

This inequality holds when b is a positive integer. To prove that the inequality holds for positive integral values of b we write it in the form

$$2b/(4b+1) > [1 - 2/(2b+3)]^{2b+\frac{1}{2}}.$$

Expanding the right hand member in ascending powers of $2/(2b+3)$ we have an alternating decreasing series and the sum of the first five terms is greater than $[1 - 2/(2b+3)]^{2b+\frac{1}{2}}$. The sum of these five terms is

$$S = (128b^4 + 64b^3 + 1648b^2 + 1208b + 1137)/24(2b+3)^4.$$

Comparing this expression with $2b/(4b+1)$ we have

$$\frac{2b}{4b+1} - S = \frac{256b^5 + 4224b^4 + 3712b^3 + 3888b^2 - 1868b - 1137}{24(4b+1)(2b+3)^4},$$

which is obviously greater than zero when $b \geqq 1$. The value

$$x_1 = [2b/(4b+1)]^{1/(4b+1)}$$

indicates the point at which V takes a minimal value.

Let us consider next some illustrations with non-symmetric distributions with a range from c to d where c and d are both positive and $c < d$.

(d). Take $f(x) = k/x$. From $k \int_c^d dx/x = 1$, $k \log(d/c) = 1$. We find $M_2 = (d^2 - c^2)/2 \log(d/c)$. Instead of equation (39) of the symmetric case, we obtain

$$V = (1/x_1^n) \int_c^d x^n f(x)\, dx - \int_{x_1}^d f(x)\, dx,$$

and $dV/dx_1 = 0$ gives $x_1^{n+1} f(x_1) = n \int_c^d x^n f(x)\, dx$. When $n = 2$, $x_1 = (d^2 - c^2)^{\frac{1}{2}}$.

This value is greater than $M_2^{\frac{1}{2}}$ and gives a minimum for V whenever d and c are so selected that

[illegible]

(e). Take $f(x) = k[x/(1-x)]$ where $c = 0$, and $d = .9$.

From $k \int_c^d [x/(1-x)] dx = 1$ we find $k = 1/1.10259$. We next find

$$x_1^{n+1}[x_1/(1-x_1)] = n \int_0^d [x^{n+1}/(1-x)]\,dx,$$

and

$$x_1^{n+2}/(1-x_1) = n \int_0^{.9} [x^{n+1}/(1-x)]\,dx.$$

When $n = 2$,

$$x_1^4/(1-x_1) = 1.509.$$

This equation has a root between .76 and .77 which corresponds to the minimal value of V.

In conclusion we may say that we have answered the question raised at the beginning of this section for a considerable number of simple functions for which the difference between the two members of the Tchebycheff inequality is a minimum at a point x_1 within the interval from $x = tM_n^{1/n}$ to the upper bound of the distribution. Although certain other functions have been found for which V takes a minimal value, it seems that the cases considered above are sufficient to illustrate the method which may be applied to a function $f(x)$ whenever it is desirable to seek such minimal values.

THE UNIVERSITY OF IOWA.

A Generalization of Eisenstein's Canonical Cubic and Associated Forms.

By Claiborne G. Latimer.

1. Eisenstein treated a canonical cubic, associated forms and reduced associated forms.* The coefficients of the canonical cubic are functions of rational integers β, γ where $(\beta + 3\gamma\rho)(\beta + 3\gamma\rho^2) = q$, ρ being an imaginary cube root of unity and q an arbitrarily chosen prime in the form $3k + 1$. He found a necessary and sufficient condition that there exist a reduced form belonging to an integer a and also found the number of such forms.

In this paper we shall consider forms similar to those mentioned above, except that instead of a prime q, we shall use an integer D, which is the product of distinct primes each in the form $3k + 1$. For these forms we shall obtain results analogous to those of Eisenstein referred to above, our results being equivalent to his when D is a prime.

The method used here, which is much shorter than Eisenstein's, involves the use of the theory of algebraic numbers. In particular, we shall use the factorable, primitive or *einheit* forms corresponding to the ideals in a cubic field of discriminant D^2 and also certain results obtained by the writer in another paper, hereafter referred to as J.† It will be found that there is a one-to-one correspondence between the ideals of a certain type in the above mentioned field and the reduced associated forms. For the case $q = 7$, Nowlan showed that Eisenstein's canonical cubic could be transformed into the norm of the general integral number of a cubic field of discriminant 7^2.‡ All the ideals in this field are principal and hence the corresponding forms are equivalent to the above mentioned norm.

2. Let $D = q_1 \cdot q_2 \cdot \cdot \cdot q_n$ where the q's are distinct primes in the form $3k + 1$. Then $q_i = \lambda_i \lambda_i'$ $(i = 1, 2, \cdots, n)$, where λ_i, λ_i' are conjugate imaginary primary primes of $F(\rho)$, and $\prod_{i-1}^{n} \lambda_i = \beta + 3\gamma\rho$, where β, γ are rational integers, $\beta \equiv (-1)^n \pmod{3}$. Consider the equation

$$x^3 - 3Dx - (2\beta - 3\gamma)D = 0. \tag{1}$$

* *Journal für Mathematik*, Vol. 28, pp. 289-374.

† "On the Prime Ideals in the General Cubic Galois Field," *American Journal of Mathematics*, Vol. 51 (1929), pp. 293-304.

$2\beta - 3\gamma$ is prime to D, since $4D = (2\beta - 3\gamma)^2 + 27\gamma^2$, and therefore (1) is irreducible. Let F be the field defined by a root of (1).* By Woronoj's method of determining a basis of the integral numbers of a cubic field,† the following numbers constitute a basis of F:

$$1, \ \theta = (-\xi + \theta_1)/\delta, \qquad \omega_1 = (\xi^2 - 3D + \xi\theta_1 + \theta_1^2)/\delta^2\sigma$$

where δ, σ, ξ are defined as follows. If each of the following conditions is satisfied, $\delta = 3$; otherwise, $\delta = 1$.

(2) $\qquad 3D \equiv 3 \pmod 9, \qquad \pm(2\beta - 3\gamma)D \equiv -3D + 1 \pmod{27}.$

σ is the largest integer such that the congruences

(3) $\quad x^3 - 3Dx - (2\beta - 3\gamma)D \equiv 0 \pmod{\delta^3\sigma^2}, \qquad 3x^2 - 3D \equiv 0 \pmod{\delta^2\sigma},$

have a common solution, ξ. The discriminant of (1) is $3^6D^2\gamma^2$ and hence the discriminant of F is $3^6D^2\gamma^2/\delta^6\sigma^2$. ($2_1$) is obviously satisfied and it will be found that (2_2) is also satisfied, the upper or lower sign being read according as n is odd or even. Hence $\delta = 3$. Then σ divides $D\gamma$. D contains no square factor and is prime to $3(2\beta - 3\gamma)$ and therefore by (3), σ is prime to D. If in (3) we set $\delta = 3$, $\sigma = \gamma$ the resulting congruences have the common solution $x = -\beta$. Hence $\delta = 3$, $\sigma = \gamma$ $\xi = -\beta$ and the following numbers constitute a basis of F.

$$1, \ \theta = (\beta + \theta_1)/3, \qquad \omega = (-\beta\theta + \theta^2)/\gamma.$$

Let θ be a root of the cubic equation $g(x) = 0$, the discriminant of which is $D^2\gamma^2$. Then by Theorem 2, J, we have the

LEMMA. *A rational prime p, not a divisor of D, is the product of three distinct conjugate prime ideals of F or is a prime of F according as*

(4) $$g(x) \equiv 0 \pmod p$$

has a solution or has no solution.

3. Let p be a prime, not a divisor of $6\gamma D$, such that $\{p\} = \mathfrak{P}\mathfrak{P}'\mathfrak{P}''$ where the $\mathfrak{P}$'s are distinct prime ideals of F. By the Lemma, (4) has a solution. Since the discriminant of g is a square, by a result due to Woronoj,‡

---

* If D is a prime, F is the field defined by one of the three periods of the imaginary Dth roots of unity. See Bachmann, *Die Lehre von der Kreisteilung*, p. 213.

† "Über die ganzen algebraischen Zahlen die von einer Wurzel der Gleichung dritten Grades abhängen." See also J, § 2.

‡ Dickson, *History of the Theory of Numbers*, Vol. 2, pp. 253-4.

it follows that (4) has three solutions, b, b', b'', no two of which are congruent, modulo p. γD is not divisible by p and therefore after adding proper multiples of p to b, b', b'' we may assume

$$g(b) = (b-\theta)(b-\theta')(b-\theta'') = p^m s, \quad g(b') = p^m s', \quad g(b'') = p^m s''$$

where m is an arbitrarily chosen positive integer and s, s', s'' are prime to p. Then $\mathfrak{P}$ divides one and only one of the above factors of $g(b)$, say $b-\theta$. For if it divided two of them, it would divide their difference and then p would divide γD, contrary to hypothesis. $\mathfrak{P}^m$ is therefore the g. c. d. of p^m and $b-\theta$; or $\mathfrak{P}^m = \{p^m, b-\theta\}$. Then

$$\mathfrak{P}^a = \{p^a, b-\theta\}, \qquad \mathfrak{P}'^a = \{p^a, b'-\theta\},$$
$$\mathfrak{P}''^a = \{p^a, b''-\theta\} \qquad (\alpha = 1, 2, \cdots, m).$$

Let $\mathfrak{A}$ be an ideal of F such that it is not divisible by a principal ideal $\{p\}$, where p is a rational prime, and such that $N(\mathfrak{A}) = a = p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdots p_k^{\lambda_k}$ where the p's are distinct primes, none of which divide $6\gamma D$. We may then assume

$$\mathfrak{A} = (\mathfrak{P}_1^{\xi_1}\mathfrak{P}_1'^{\eta_1})(\mathfrak{P}_2^{\xi_2}\mathfrak{P}_2'^{\eta_2}) \cdots (\mathfrak{P}_k^{\xi_k}\mathfrak{P}_k'^{\eta_k})$$

where $\mathfrak{P}_i \neq \mathfrak{P}_i'$, $N(\mathfrak{P}_i) = N(\mathfrak{P}_i') = p_i$, $\xi_i + \eta_i = \lambda_i$ $(i = 1, 2, \cdots, k)$. Let $\mathfrak{A}_1$ be an ideal obtained from $\mathfrak{A}$ by replacing every product $\mathfrak{P}_i^{\xi_i}\mathfrak{P}_i'^{\eta_i}$ by one of its conjugates. If an ideal $\mathfrak{G}$ is the product of two ideals $\mathfrak{A}$ and $\mathfrak{A}_1$ as above defined, we shall say that $\mathfrak{G}$ is of type $P(a)$. Let $\mathfrak{P}^\xi\mathfrak{P}'^\eta$ be one of the factors of $\mathfrak{A}$ displayed above and let $\mathfrak{P}'^\xi\mathfrak{P}''^\eta$ be the corresponding factor of $\mathfrak{A}_1$. Assuming that $\xi \geqq \eta$, consider the product,

$$(\mathfrak{P}^\xi\mathfrak{P}'^\eta)(\mathfrak{P}'^\xi\mathfrak{P}''^\eta) = \{p^\eta\}\mathfrak{P}^{\xi-\eta}\mathfrak{P}'^\xi = \{p^\eta,\}\{p^{\xi-\eta}, b-\theta\}\{p^\xi, b'-\theta\},$$

where b, b' are rational integers, incongruent modulo p, determined as above with $m = \xi$. It may be shown that the rational integers in this product are identical with the multiples of $p^{\xi+\eta} = p^\lambda$. If $\eta > \xi$ we obtain the same result in a similar manner. By repetition of this argument we find that the rational integers in $\mathfrak{A}\mathfrak{A}_1$ are identical with the multiples of a. Therefore $\mathfrak{A}\mathfrak{A}_1$ has a canonical basis

$$\mathfrak{A}\mathfrak{A}_1 = [a, g + s\theta, h + k\theta + t\omega]$$

where $g, \cdots, t$ are rational integers. We may assume $0 \leqq h < a$, $0 \leqq g$, [illegible]

It is known [illegible] that if β_1, β_2, β_3 are integers of F which constitute a basis

[illegible]

of an ideal $\mathfrak{B}$, then the factorable form

$$U = N(\beta_1 x + \beta_2 y + \beta_3 z)/N(\mathfrak{B})$$

is primitive, i. e. the coefficients are rational integers, the g. c. d. of which is unity. We shall consider later the form corresponding to $\mathfrak{A}\mathfrak{A}_1$; namely,

$$G(x, y, z) = (1/a^2)N[ax + (g + s\theta)y + (p + k\theta + t\omega)z]$$

4. Consider the form

$$\Phi(u, v, w) = u^3 + D(\beta + 3\gamma\rho)(v + w\rho)^3 + D(\beta + 3\gamma\rho^2)(v + w\rho^2) - 3Du(v + w\rho)(v + w\rho^2).$$

A ternary cubic form $f = ax^3 + \cdots$ where $a > 0$, will be said to be *associated with* Φ if its coefficients are rational integers, the g. c. d. of which is unity, and if Φ is transformed into $a^2 f$ by

$$S \qquad \begin{aligned} u &= ax + by + b'z \\ v &= \phantom{ax +{}} cy + c'z \\ w &= \phantom{ax +{}} dy + d'z \end{aligned}$$

where $b, \cdots, d'$ are rational integers such that $cd' - c'd = a$. Such a form shall also be said to *belong to* a. If $c' = 0$; $0 \leqq b$, $b' < a$; $0 \leqq d < d'$; f will be said to be a *reduced associated form.** To Φ apply the transformation

$$T_1 \qquad \begin{aligned} u &= \alpha_0 \phantom{{}+\alpha_1} + 2D\alpha_2 \\ v &= \phantom{\alpha_0 + \alpha_1 +{}} 3\gamma\alpha_2 \\ w &= \phantom{\alpha_0 +{}} \alpha_1 + \beta\alpha_2. \end{aligned}$$

The resulting form will be found to be $\psi(\alpha_0, \alpha_1, \alpha_2) = N(\alpha_0 + \alpha_1\theta_1 + \alpha_2\theta_1^2)$, where θ_1 is a root of (1).† Setting $\alpha_0 + \alpha_1\theta_1 + \alpha_2\theta_1^2 = x + y\theta + z\omega$, we find that ψ is transformed into $\phi(x, y, z) = N(x + y\theta + z\omega)$ by

$$T_2 \qquad \begin{aligned} \alpha_0 &= x + (1/3)\beta y - (2\beta^2/9\gamma)z \\ \alpha_1 &= \phantom{x +{}} (1/3)y - (\beta/9\gamma)z \\ \alpha_2 &= - (1/9\gamma)z. \end{aligned}$$

Consider the form G, defined in § 3. ϕ is transformed into $a^2 G$ by the transformation T_3 with the matrix

* If D is a prime, our definitions of Φ and *reduced associated forms* are identical with Eisenstein's. Our definition of an *associated form* is equivalent to his. See *ibid.*, pp. 297, 333, 334, 318, 296.

† ψ may be obtained from (6), J, by replacing a, b by $-3D$ and $(3\gamma - 2\beta)D$ respectively.

$$\begin{pmatrix} a & g & h \\ 0 & s & k \\ 0 & 0 & t \end{pmatrix}.$$

Apply to G the transformation T_4 with the matrix,

$$\begin{pmatrix} 1 & \xi & \eta \\ 0 & 0 & 3 \\ 0 & 3 & 3\zeta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & \xi & \eta \\ 0 & 1 & \zeta \\ 0 & 0 & 1 \end{pmatrix}$$

where ξ, η, ζ are arbitrarily chosen rational integers, and let f be the resulting form. G is primitive and T_4 is the product of two transformations as indicated above. The form obtained by the first transformation is primitive since a, the leading coefficient of G, is prime to 3. The second transformation is unitary and a is odd. It may then be shown that f is primitive.*

Φ is transformed into $a^2 f$ by $T_1 T_2 T_3 T_4$. The matrix of this transformation is

$$\begin{pmatrix} a & a\xi + \mu & a\eta + \nu \\ 0 & t & \zeta t \\ 0 & k & s + k\zeta \end{pmatrix}$$

where μ, ν are certain integers independent of ξ, η. By § 3, $st = a$. Hence f is an associated form belonging to a. If we choose ξ, η, ζ so that $0 \leqq a\xi + \mu < a$, $0 \leqq a\eta + \nu < a$, $\zeta = 0$, then f is a reduced associated form.

It follows that a sufficient condition for the existence of a reduced associated form belonging to a, where a is prime to $6\gamma D$, is that there exist an ideal of type $P(a)$; or that every prime ideal factor of a be of the first degree. In the next paragraph we shall show that this condition is also necessary.

5. Let f be a reduced associated form belonging to an integer $a = p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdots p_k^{\lambda_k}$ where the p's are distinct primes, none of which divide $6\gamma D$. Then ϕ is transformed into $a^2 f$ by $T_2^{-1} T_1^{-1} S$, where in S, $c' = 0$. The matrix of this transformation is

$$\begin{pmatrix} a & \delta & b' - d'\beta \\ 0 & 3d & 3d' \\ 0 & 3c & 0 \end{pmatrix},$$

where $\delta = b + 2(\beta - 2\gamma)c - d\beta$. Then

(3) $$f = N[ax + (\delta + 3d\theta + 3c\omega)y + (b' - d'\beta + 3d'\theta)z]/a^2.$$

Let $\mathfrak{B}$ be the g. c. d. of the above coefficients; or

* [illegible]

$$\mathfrak{B} = \{a,\ \delta + 3d\theta + 3c\omega,\ b' - d'\beta + 3d'\theta\}.$$

The form $a^2f/N(\mathfrak{B})$ has rational integral coefficients, the g. c. d. of which is unity.* The same is true of f and hence $N(\mathfrak{B}) = a^2$. $a = cd'$ is prime to 3 and therefore we may determine integers λ, ϵ such that $\lambda ad' + 3\epsilon = 1$. Then,

$$a, \qquad \epsilon(\delta + 3d\theta) + c\omega, \qquad \epsilon(b' - d'\beta) + d'\theta$$

belong to $\mathfrak{B}$. These numbers constitute a basis of $\mathfrak{B}$ since $N(\mathfrak{B}) = acd'$; or

$$\mathfrak{B} = [a,\ \ \epsilon(b' - d'\beta) + d'\theta, \qquad \epsilon(\delta + 3d\theta) + c\omega].$$

Consider the factorable, primitive corresponding to $\mathfrak{B}$; namely

$$f'(x', y', z') = (1/a^2)N\{ax' + [\epsilon(b' - d'\beta) + d'\theta]y' + [\epsilon(\delta + 3d\theta) + c\omega]z'\}.$$

Equating the expression in brackets to the corresponding expression in (3) we find that f' is transformed into f by

$$\begin{aligned} x' &= x + \kappa y + \tau z \\ y' &= 3\lambda ad\, y + 3z \\ z' &= 3y \end{aligned}$$

where κ, τ are rational integers. The determinant of this transformation is -9. We shall now show that $\mathfrak{B}$ is of type $P(a)$.

Suppose $\mathfrak{B}$ were divisible by a prime ideal of the third degree, $\{p\}$. Let $\mathfrak{B} = \{p^n\}\mathfrak{B}_1$, where $\mathfrak{B}_1$ and hence $N(\mathfrak{B}_1)$ is prime to p. By the above basis of $\mathfrak{B}$, c and d' are divisible by p^n. Then a^2 is divisible by p^{4n}, whereas $a^2 = N(\mathfrak{B}) = p^{3n}N(\mathfrak{B}_1)$ where $N(\mathfrak{B}_1)$ is prime to p. We may then assume

$$\mathfrak{B} = (\mathfrak{P}_1^{\xi_1}\mathfrak{P}_1'^{\zeta_1}\mathfrak{P}_1''^{\eta_1})(\mathfrak{P}_2^{\xi_2}\mathfrak{P}_2'^{\zeta_2}\mathfrak{P}_2''^{\eta_2})\cdots(\mathfrak{P}_k^{\xi_k}\mathfrak{P}_k'^{\zeta_k}\mathfrak{P}_k''^{\eta_k})$$

where $\mathfrak{P}_i$, $\mathfrak{P}_i'$, $\mathfrak{P}_i''$ are distinct prime ideals of norm p_i, $\xi_i + \eta_i + \zeta_i = 2\lambda_i$, $\zeta_i \geqq \xi_i \geqq \eta_i \geqq 0$ $(i = 1, 2, \cdots, k)$. Consider the product $\mathfrak{P}_1^{\xi_1}\mathfrak{P}_1'^{\zeta_1}\mathfrak{P}_1''^{\eta_1}$. Dropping subscripts, we have

$$\mathfrak{P}^{\xi}\mathfrak{P}'^{\zeta}\mathfrak{P}''^{\eta} = \{p^{\eta}\}\mathfrak{M},\ \mathfrak{M} = \mathfrak{P}^{\xi-\eta}\mathfrak{P}'^{\zeta-\eta} = \{p^{\xi-\eta}, b - \theta\}\{p^{\zeta-\eta}, b' - \theta\}$$

where b, b' are rational integers, determined as in § 2 with $m = 2\lambda$. The rational integers in $\mathfrak{B}$ are identical with the multiples of a and hence those in $\mathfrak{M}$ are the multiples of $p^{\lambda-\eta}$. Since $N(\mathfrak{M}) = p^{2\lambda-3\eta}$, it follows that $\mathfrak{M}$ has the canonical basis

$$\mathfrak{M} = [p^{\lambda-\eta},\ R + S\theta,\ T + U\theta + V\omega]$$

---

* Bachmann, *Allgemeine Arithmetik der Zahlenkörper*, pp. 425-6.

where $R, \cdots, V$ are positive integers and $SV = p^{\lambda-2\eta}$. $\mathfrak{M}$ contains $(b-\theta)(b-\theta') = bb' + (\beta - b - b')\theta + \gamma\omega$ and since a is prime to γ it follows that $V = 1$. $\mathfrak{M}$ contains $p^{\zeta-\eta}(b-b')$ and hence it contains $p^{\zeta-\eta}$. Therefore $\lambda \leqq \zeta$, or $\lambda \geqq \xi + \eta$. $\mathfrak{M}$ also contains $p^{\xi-\eta}(b'-\theta)$ and therefore $S = p^{\lambda-2\eta}$ divides $p^{\xi-\eta}$; or $\lambda \leqq \xi + \eta$. Hence $\lambda = \xi + \eta = \zeta$ and

$$\mathfrak{P}^{\xi}\mathfrak{P}'^{\zeta}\mathfrak{P}''^{\eta} = (\mathfrak{P}^{\xi}\mathfrak{P}'^{\eta})(\mathfrak{P}'^{\xi}\mathfrak{P}''^{\eta})$$

Repeating this argument, we find that $\mathfrak{B}$ is of type $P(a)$.

If $\alpha_1, \alpha_2, \alpha_3$ and $\beta_1, \beta_2, \beta_3$ are two bases of an ideal $\mathfrak{B}$, then $\alpha_i = \sum_j c_{ij}\beta_j$ $(i = 1, 2, 3)$ where c_{ij} are rational integers and the determinant $|c_{ij}| = \pm 1$. It follows that the primitive factorable forms corresponding to these bases are equivalent. We have therefore

THEOREM 1. *Let F be the cubic field defined by a root of* (1) *and let $\mathfrak{G}$ be an ideal of F of type $P(a)$, where a is prime to $6\gamma D$.* If $\alpha_1, \alpha_2, \alpha_3$ *are integers of F which constitute a basis of $\mathfrak{G}$, then the primitive, factorable form*

$$(1/a^2)N(\alpha_1 x + \alpha_2 y + \alpha_3 z)$$

may be transformed into a reduced associated form belonging to a by a transformation, T, with rational integral coefficients, of determinant ± 9.

Conversely, if f is such a reduced associated form, where a is prime to $6\gamma D$, there exists an ideal of type $P(a)$ such that every primitive factorable form corresponding to this ideal may be transformed into f by such a transformation T.

COROLLARY. *If a is a positive integer, prime to $6\gamma D$, there exists a reduced associated form belonging to a if and only if every rational prime factor of a is the norm of an ideal of F.**

By the corollary and Theorem 4, J, we have

THEOREM 2. *If a is a positive integer, prime to $6\gamma D$, a sufficient condition that there exist a reduced associated form belonging to a is that every prime factor of a be a cubic residue of D. If D is a prime, this condition is also necessary.*†

6. In this paragraph we shall show that there is a one-to-one correspondence between the ideals of type $P(a)$ and the reduced associated forms

* Theorem 3, J, gives a necessary and sufficient condition that a rational prime

belonging to a. We shall then find the number of such ideals and hence the number of such forms.

In the preceding paragraph we saw that to every reduced form there corresponded an ideal $\mathfrak{B}$ of type $P(a)$. Let f_1, f_2 be two reduced forms belonging to a and suppose they correspond to the same ideal $\mathfrak{B}$. Then Φ is transformed into a^2f_1, a^2f_2 by transformations with the following matrices:

$$\begin{pmatrix} a & b_1 & b_1' \\ 0 & c_1 & 0 \\ 0 & d_1 & d_1' \end{pmatrix}, \quad \begin{pmatrix} a & b_2 & b_2' \\ 0 & c_2 & 0 \\ 0 & d_2 & d_2' \end{pmatrix},$$

where $0 \leqq b_i, b_i' < a$; $0 \leqq d_i < d_i'$; $c_i d_i' = a$ $(i = 1, 2)$. We have

$$\begin{aligned} \mathfrak{B} &= [a, \epsilon_1(b_1' - \beta d_1') + d_1'\theta, \epsilon_1(\delta_1 + 3d_1\theta) + c_1\omega] \\ &= [a, \epsilon_2(b_2' - \beta d_2') + d_2'\theta, \epsilon_2(\delta_2 + 3d_2\theta) + c_2\omega] \end{aligned}$$

where the ϵ_i are determined by $\lambda_i a d_i' + 3\epsilon_i = 1$ and $\delta_i = b_i + 2(\beta - 3\gamma)c_i - d_i\beta$. From the above sets of basal numbers it follows that $c_1 = c_2$, $d_1' = d_2'$ and hence we may assume $\epsilon_1 = \epsilon_2$. Since the numbers of each set may expressed in terms of the other, it is readily shown that the forms f_1 and f_2 are identical. Hence the number of reduced forms belonging to a is at most equal to the number of ideals of type $P(a)$.

From every ideal of type $P(a)$ we obtain a reduced form belonging to a. Suppose we obtain the same form f from two ideals. Let G and G' be the corresponding primitive, factorable forms, as defined at the end of § 3. Then G and G' are transformed into f by transformations T_4 and T_4' respectively, with matrices

$$\begin{pmatrix} 1 & \xi & \eta \\ 0 & 0 & 3 \\ 0 & 3 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & \xi' & \eta' \\ 0 & 0 & 3 \\ 0 & 3 & 0 \end{pmatrix}.$$

Then G' is transformed into G by a transformation with the matrix

$$\begin{pmatrix} 1 & (\eta' - \eta)/3 & (\xi' - \xi)/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The coefficients of G and G' are rational integers and hence the same is true of the coefficients of this transformation. It follows that the ideals corresponding to G and G' are equal. Hence the number of reduced forms belonging to a is at least equal to the number of ideals of type $P(a)$. Therefore the number of such ideals is exactly equal to the number of reduced forms.

Consider the number of ideals of type $P(a)$, assuming that there is one. If $a = p^\lambda$, p a prime, the distinct ideals of type $P(a)$ are

$$(\mathfrak{P}^{\lambda-i}\mathfrak{P}'^{i})(\mathfrak{P}'^{\lambda-i}\mathfrak{P}''^{i}), \qquad (i = 1, 2, \cdots, \lambda).$$

where the $\mathfrak{P}$'s are distinct primes of norm p, and the 2λ ideals obtained from these by replacing $\mathfrak{P}$, $\mathfrak{P}'$, $\mathfrak{P}''$ by $\mathfrak{P}'$, $\mathfrak{P}''$, $\mathfrak{P}$ and by $\mathfrak{P}''$, $\mathfrak{P}$, $\mathfrak{P}'$ respectively. Hence there are exactly 3λ such ideals. If $a = p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdots p_k^{\lambda_k}$, where the p's are distinct primes, we obtain $3^k \lambda_1 \cdot \lambda_2 \cdots \lambda_k$ ideals of the required type.

THEOREM 3. *If $a = p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdots p_k^{\lambda_k}$, where the p's are distinct primes, none of which divide $6\gamma D$, and if there is one reduced form belonging to a, there are exactly*

$$3^k \lambda_1 \cdot \lambda_2 \cdots \lambda_k$$

*such reduced forms.**

7. In § 2 we defined $\beta + 3\gamma\rho$ as a product of primary primes. We may obtain $2^{n-1} = t$ such products, $\beta_i + 3\gamma_i\rho$ $(i = 1, 2, \cdots, t)$ such that the norm of each is D and no two of them are conjugates. For each of these products we obtain as in § 2 a field of discriminant D^2 and also a corresponding Φ_i as defined in § 4. The Lemma and Theorems 1, 2, 3 may be applied to each of these Φ_i and the corresponding field F_i $(i = 1, 2, \cdots, t)$. We shall now prove.

THEOREM 4. *If a is a positive integer prime to $6\gamma_1 \cdot \gamma_2 \cdots \gamma_t \cdot D$, then for every form Φ_i $(i = 1, 2, \cdots, t)$ there exists a reduced associated form belonging to a if and only if a is a cubic residue of D.*

In view of Theorem 2, it suffices to prove the necessary condition when $n > 1$. Suppose a is an integer, prime to $6\gamma_1 \cdot \gamma_2 \cdots \gamma_t \cdot D$, such that for every Φ_i there is a reduced associated form belonging to a, and suppose p is a prime factor of a. By the corollary of Theorem 1, p is the product of three conjugate ideals in each of the fields F_i $(i = 1, 2, \cdots, t)$. Let $\beta_1 + 3\gamma_1\rho = \prod_{i=1}^{n} \lambda_i$, $\beta_2 + 3\gamma_2\rho = \lambda_1' \prod_{i=2}^{n} \lambda_i$ where the λ_i are primary primes of $F(\rho)$ and λ_1' is the conjugate of λ_1. Let F_1 and F_2 be the corresponding fields, and let $p = \pi \cdot \pi'$ or $p = \pi$ where π is a primary prime of $F(\rho)$. Since p is a composite in F_1 and in F_2 by Theorem 3, I, and (c) of Lemma 2, I, we have

$$\prod_{i=1}^{n} \left[\frac{\pi}{\lambda_i}\right] = \prod_{i=1}^{n} \left[\frac{\pi}{\lambda_i'}\right], \quad \left[\frac{\pi}{\lambda_1'}\right] \prod_{i=2}^{n} \left[\frac{\pi}{\lambda_i}\right] = \left[\frac{\pi}{\lambda_1}\right] \prod_{i=2}^{n} \left[\frac{\pi}{\lambda_i'}\right],$$

* [illegible]

where the brackets are Eisenstein's symbols for the cubic character of π with respect to the corresponding primes. Therefore

$$\left[\frac{\pi}{\lambda_1}\right] = \left[\frac{\pi}{\lambda_1'}\right].$$

Squaring both sides, we have

$$\left[\frac{\pi'}{\lambda_1'}\right] = \left[\frac{\pi}{\lambda_1}\right].$$

If $p = \pi \cdot \pi'$ by multiplying these equations we obtain

$$\left[\frac{p}{\lambda_1}\right] = \left[\frac{p}{\lambda_1'}\right],$$

which is also valid if $p = \pi$. But

$$\left[\frac{p}{\lambda_1}\right]^2 = \left[\frac{p}{\lambda_1'}\right].$$

Hence each of these symbols is unity; or p is a cubic residue of each of the non-associated prime factors of $\lambda_1\lambda_1' = q_1$. Repetition of this argument shows that p is a cubic residue of $q_1 q_2 \cdot \cdot \cdot q_n = D$. This proves the theorem.

UNIVERSITY OF KENTUCKY.

Quadruples of Space Curves.*

By A. R. Jerbert.

1. *Introduction.* Several papers have appeared recently which employ the notion of configurations whose elements are in one-to-one-correspondence. These papers are: "Ruled Surfaces with their Generators in one-to-one Correspondence," by E. P. Lane; † "Triads of Ruled Surfaces," by A. F. Carpenter,‡ and "Triads of Plane Curves," by the writer.§

The present paper is a logical continuation of the last one cited. It bears, however, a certain relation to Professor Lane's paper. This consists in the fact that in both papers the analytic basis is a system of four, linear, homogeneous, differential equations of the first order. In each case the given system of equations defines, to within a projectivity, four space curves.

From this point on the viewpoints are divergent. In Professor Lane's paper the four points are paired. One pair defines the ruling of one surface, while the other pair defines the corresponding ruling for a second surface. The curves serve merely as directrix curves for the surfaces and as such they are replaced, at an early stage, by more suitable curves, namely, the "intersector curves." ¶

In the present paper the curves given by the defining system are regarded as the entity under consideration. To this end, only transformations which preserve the identity of the curves are permitted. These transformations are, the change of parameter and the transformation which multiplies the homogeneous coördinates by an arbitrary factor.

2. *Defining System.* Let us suppose that the four homogeneous coördinates of a point x_1 are given as functions of a parameter t; that is, $x_1^{(i)} = f_i(t)$, $(i = 1, 2, 3, 4)$. As the parameter t varies, the point x_1 describes a curve C_{x_1}. In the same fashion we shall assume that three other points, x_2, x_3 and x_4, have their coördinates given as functions of the same parameter t.

As t varies, the four points describe the space curves C_{x_1}, C_{x_2}, C_{x_3}, and C_{x_4} which are in point correspondence by virtue of their expression in terms of

* Read before the American Mathematical Society [illegible]ne 2, 1928.

† *Transactions of the American Mathematical Society*, Vol. 25 (1923), p. 281.

‡ *Transactions of the American Mathematical Society*, Vol. 29 (1927), p. 254.

§ [illegible]

¶ [illegible]

the common parameter t. All the functions involved will be assumed to be continuous and differentiable except at most for isolated values of the parameter.

We proceed to find for the four curves a defining system of the form,

$$x_i' = \sum_{j=1}^{4} a_{ij}\, x_j \, (i = 1, 2, 3, 4), \tag{1}$$

where differentiation is with respect to t. The problem reduces to computing the coefficients a_{ij}. To solve for the a_{1i} $(i = 1, 2, 3, 4)$ substitute in the first of equations (1), in turn, the first, second, third, and fourth coördinates of the points x_1, x_2, x_3, x_4. The resulting system of four equations can be solved for the a_{1i} providing that the determinant

$$\Delta \equiv \mid x_1^{(1)}, x_2^{(2)}, x_3^{(3)}, x_4^{(4)} \mid \tag{2}$$

is not identically zero.

In the same manner the remaining coefficients can be found provided always that the determinant (2) does not vanish identically. This assumption we shall make. Geometrically it amounts to asserting that the correspondence between the curves is such that the corresponding points are not coplanar, except at most, for isolated values of t.

If in place of the four points $x_1, \cdots x_4$ we employ four new points obtained from the latter by the most general projectivity we find that they give rise to the same defining system (1). System (1), therefore, defines the configuration of four associated space curves only to within a projectivity and is accordingly suitable for the investigation of the projective differential properties of such a configuration.

3. *Canonical System.* As stated in the introduction, the object of this paper is to study the mutual relation of the curves defined by (1). With this in view the only permissible transformations are the change of parameter and the multiplicative transformation on the homogeneous coördinates. We shall employ the latter to obtain a canonical form for system (1). To this end let us put

$$x_i = \lambda_i \bar{x}_i \; (i = 1, 2, 3, 4), \tag{3}$$

where the λ_i $(i = 1, 2, 3, 4)$, are arbitrary functions of the parameter t. If we substitute these expressions in (1) we obtain after transposition and division,

$$\bar{x}_i' = \sum_{j=1}^{4} (a_{ij}\lambda_j - \delta_{ij}\lambda_j')\bar{x}_j/\lambda_i \; (i = 1, 2, 3, 4). \tag{4}$$

From (4) it is apparent that we can choose the arbitrary proportionality

factors in such a way as to render the coefficients in the principal diagonal equal to zero. For this purpose we shall take

$$\lambda_i = k_i e \exp \int a_{ii} \delta t, \qquad (i = 1, 2, 3, 4).$$

We obtain, accordingly, the defining system

$$
\begin{aligned}
\bar{x}_1' &= \qquad\qquad\ \bar{a}_{12}\,\bar{x}_2 + \bar{a}_{13}\,\bar{x}_3 + \bar{a}_{14}\,\bar{x}_4, \\
\bar{x}_2' &= \bar{a}_{21}\,\bar{x}_1 + \qquad\quad + \bar{a}_{23}\,\bar{x}_3 + \bar{a}_{34}\,\bar{x}_4, \\
\bar{x}_3' &= \bar{a}_{31}\,\bar{x}_1 + \bar{a}_{32}\,\bar{x}_2 \qquad\qquad + \bar{a}_{34}\,\bar{x}_4, \\
\bar{x}_4' &= \bar{a}_{41}\,\bar{x}_1 + \bar{a}_{42}\,\bar{x}_2 + \bar{a}_{43}\,\bar{x}_3 \qquad\qquad ,
\end{aligned}
\tag{5}
$$

where

$$\bar{a}_{ij} = a_{ij}\,\lambda_j/\lambda_i. \tag{6}$$

One other "permissible" transformation remains and that is the change of parameter. It is readily verified that the effect of such a transformation is to multiply the coefficients in the defining system by a factor. Since we can not utilize the parameter transformation to any advantage, we shall pay no further heed to it. For the purposes of this paper, therefore, defining system (5) will serve as a canonical system.

4. *Invariants and Covariants.* From (6) we see that such products of the coefficients as $a_{ij}\,a_{ji}$ are invariants under the λ-transformation. The coefficient a_{ij} carries with it a factor λ_j/λ_i whereas a_{ji} involves the factor λ_i/λ_j. The invariancy of $a_{ij}\,a_{ji}$ is rendered apparent if we think of an inner subscript as cancelling a like outer subscript. With this in mind, such a product as $a_{23}\,a_{34}\,a_{42}$ is obviously invariant.

As a corollary to the last remark we note that if in the determinant of the coefficients a_{ij} (equation 5) we replace the diagonal terms by 1's and expand by the ordinary rule all the terms in the expansion are invariants.

With respect to covariants we note similarly that such an expression as $a_{41}\,x_3 - a_{31}\,x_4$ carries with it, on being transformed, a factor $\lambda_1/\lambda_3\,\lambda_4$ and is therefore a relative covariant. The following invariants are listed and named for convenience of reference.

$$
\begin{aligned}
&A_{12} = a_{12}\,a_{21},\ A_{13} = a_{13}\,a_{31},\ A_{14} = a_{14}\,a_{41}, \\
&A_{23} = a_{23}\,a_{32},\ A_{24} = a_{24}\,a_{42},\ A_{34} = a_{34}\,a_{43}, \\
&M_1 = a_{23}\,a_{34}\,a_{42}, \qquad\qquad N_1 = a_{32}\,a_{43}\,a_{24}, \\
&M_2 = a_{13}\,a_{34}\,a_{41}, \qquad\qquad N_2 = a_{31}\,a_{14}\,a_{43}, \\
&M_3 = a_{21}\,a_{14}\,a_{42}, \qquad\qquad N_3 = a_{12}\,a_{24}\,a_{41}, \\
&M_4 = a_{12}\,a_{23}\,a_{31}, \qquad\qquad N_4 = a_{21}\,a_{13}\,a_{32},
\end{aligned}
\tag{7}
$$

We can characterize the canonical system (5) geometrically if we recall

The canonical system implies that we have chosen the proportionality factors (the λ_i) in such a fashion that x_1' is the point where the tangent to C_{x_1} meets the opposite face of the tetrahedron. Similar statements hold for each of the vertices of the tetrahedron.

Since the work that follows will be based entirely upon the canonical system (5), no ambiguity will occur if we write the coefficients which appear in (5) without the bars.

5. *Dual Considerations.* It will be recalled that the assumption was made at the outset, that the correspondence between the four curves was to be such that corresponding points should not, in general, be coplanar. It follows that if we join corresponding points by straight lines a non-degenerate tetrahedron is formed. Since the parameter t can take a singly infinite number of values we see that the study of the four associated curves is equivalent to the study of a one parameter family of tetrahedrons.

Since a tetrahedron is determined by its face planes as well as by its vertices it will prove useful to find a defining system for the coördinates of these planes. If we denote by u_i the plane opposite the vertex x_i it follows that the coördinates of u_i, i. e. $u_i^{(k)}$ $(k = 1, 2, 3, 4)$, are proportional to the cofactors of $x_i^{(k)}$ $(k = 1, 2, 3, 4)$, in the determinant Δ (equation 2). We shall assume for the coördinates of u_i the cofactors themselves. By direct computation we find by the aid of equations (5) the defining system

$$
(8) \qquad \begin{array}{rllll}
u_1' = & & -a_{21}u_2 & -a_{31}u_3 & -a_{41}u_4, \\
u_2' = & -a_{12}u_1 & & -a_{32}u_2 & -a_{42}u_4, \\
u_3' = & -a_{13}u_1 & -a_{23}u_2 & & -a_{43}u_4, \\
u_4' = & -a_{14}u_1 & -a_{24}u_2 & -a_{34}u_3 & .
\end{array}
$$

It will be observed that the defining system (8) like (5) is in canonical form. If the computation for (8) had been based upon the uncanonical system (1) we would have obtained in place of (8) a system like it but with the diagonal terms restored.

The geometrical interpretation of equations (8) is dual to that of equations (5). Equation $(8)_1$ states that the plane u_1' passes through the intersection of the planes u_2, u_3, and u_4. Similar statements hold for equations $(8)_2$, $(8)_3$, and $(8)_4$.

6. *Characteristic Lines.* As the parameter t varies each face of the tetrahedron gives rise to a one parameter family of planes which envelop an edge of regression.

If, in any one of the face planes, we wish to determine the line which is tangent to the edge of regression, in other words the characteristic line, we find that this is most readily accomplished by the aid of equations (8).

For this purpose we have only to recall that the characteristic line represents the intersection of the plane with its immediate successor. In the dual case we determine the join of the point x_1 and its immediate successor as the line joining x_1 and x_1'. We conclude, therefore, that the characteristic line in the plane u_i may be found as the intersection of the planes u_i and u_i'.

It proves convenient at this point to introduce a local coördinate system. We are, so to speak, already provided with a suitable reference tetrahedron. There remains merely the choice of unit point. We shall choose the latter in such a manner that an expression of the form $y_i\, x_i$ $(i = 1, 2, 3, 4)$ shall represent the point whose local coördinates are y_i $(i = 1, 2, 3, 4)$.

With this choice of coördinate system the planes u_1 and u_1' become in point coördinates, the planes

$$y_1 = 0, \; a_{21}\, y_2 + a_{31}\, y_3 + a_{41}\, y_4 = 0. \tag{9}$$

Equations (9) taken simultaneously, define the characteristic line of the plane u_1 $(y_1 = 0)$. The characteristic line so defined, intersects the edges of the tetrahedron lying in its plane in three points given by the expressions

$$a_{31}\, x_2 - a_{21}\, x_3, \; a_{41}\, x_3 - a_{31}\, x_4, \; a_{41}\, x_2 - a_{21}\, x_4. \tag{10}$$

In a similar way each characteristic line determines a set of three points, twelve in all, two for each edge of the tetrahedron. The points (10) are manifestly covariant as we should be led to expect from their geometrical character. Before proceeding with the properties of the characteristic points we shall find it advisable to introduce other considerations.

7. *Tangent Planes.* Each edge $x_i\, x_j$ of the tetrahedron generates a ruled surface as the parameter t varies and the points x_i, x_j describe their respective curves. The plane determined by the points x_i, x_j, and $(x_i + \lambda x_j)'$, is tangent to the ruled surface $(x_i\, x_j)$ at the point $x_i + \lambda x_j$. For the six ruled surfaces we obtain the following tangent planes:

$$\begin{aligned}
&(x_1\, x_2), \; (a_{13} + \lambda a_{23})\, y_4 - (a_{14} + \lambda a_{24})\, y_3 = 0,\\
&(x_1\, x_3), \; (a_{12} + \lambda a_{32})\, y_4 - (a_{14} + \lambda a_{34})\, y_2 = 0,\\
&(x_1\, x_4), \; (a_{13} + \lambda a_{43})\, y_2 - (a_{12} + \lambda a_{42})\, y_3 = 0,\\
&(x_2\, x_3), \; (a_{21} + \lambda a_{31})\, y_4 - (a_{24} + \lambda a_{34})\, y_1 = 0,\\
&(x\, x), \; (a\; + \lambda a\;)\, y \; [illegible] \; (a\; + \lambda a\;)\, y = 0,\\
&(x_3\, x_4), \; (a_{31} + \lambda a_{41})\, y_2 - (a_{32} + \lambda a_{42})\, y_1 = 0.
\end{aligned} \tag{11}$$

$(11)_4$ reduces to $y_1 = 0$, if we take for λ the value $-a_{21}/a_{31}$. In other words, the point (10), is the point on the ruled surface $(x\; x)$ at which the

lar fashion we find the points on $(x_3 x_4)$ and $(x_4 x_2)$ at which the tangent planes reduce to $x_1 = 0$. They prove to be the other characteristic points given by (10).

The fact thus disclosed can be stated for a single ruled surface.

THEOREM 1. *If a plane be passed through each ruling of a ruled surface there results a one parameter family of planes which possesses the following property. Each plane is tangent to the ruled surface at that point, of the ruling through which it is passed, in which the latter is met by the characteristic line of the plane in question.*

On the other hand, if we wish to find the equation of the tangent plane to $x_1 x_2$ at x_1 we have merely to set λ in (11) equal to zero, thus obtaining the plane

$$a_{13} y_4 - a_{14} y_3 = 0. \tag{12}$$

The plane (12 meets the opposite edge $x_3 x_4$ in a point whose expression is

$$a_{13} x_3 + a_{14} x_4. \tag{13}$$

In the same manner we determine the points in which the tangent planes to $x_1 x_4$ and $x_1 x_2$ at x_1 meet the opposite edges. We obtain in this way three points in all. They are

$$a_{13} x_3 + a_{14} x_4, \; a_{12} x_2 + a_{13} x_3, \; a_{12} x_2 + a_{14} x_4. \tag{14}$$

From the manner in which the expressions (14) are formed it is clear that each of the corresponding points is collinear with the point x_1' (equation 5), and the opposite vertex. We state this result as

THEOREM 2. *In a one parameter family of tetrahedrons the planes constructed at any vertex, tangent to the ruled surfaces intersecting in the curve generated by this vertex, intersect the edges of the opposite face in a set of three points. The lines joining these points to the opposite vertices in this face are concurrent in the point at which the tangent to the curve described by the vertex in question meets the opposite face.*

8. *Space Quadrilaterals.* Under (14) we note that there are 12 such points—two on each edge of the tetrahedron. A number of theorems can be stated which are concerned with the relations that may exist between the latter set of points and those we have denoted as characteristic points. The following relations may be readily verified by reference to equations (5), (9), and two others similar to (9) for the x_4 characteristic.

$$
(15)\quad
\begin{array}{llll}
A_{12} \quad + A_{13} & = 0, \ (1^3\,4\ ), & x_1\,x_4\,x_1' \text{ meets } x_2\,x_3 & \text{in } x_1 \text{ char,} \\
a_{12}\,a_{24} + a_{13}\,a_{34} & = 0, \ (1^2\,4^2), & x_1\,x_4\,x_1' \quad `` \quad x_2\,x_3 & `` \quad x_4 \quad `` \ , \\
A_{24} \quad + A_{34} & = 0, \ (1\ \,4^3), & x_1\,x_4\,x_4' \quad `` \quad x_2\,x_3 & `` \quad x_4 \quad `` \ , \\
a_{42}\,a_{21} + a_{31}\,a_{43} & = 0, \ (4^2\,1^2), & x_1\,x_4\,x_4' \quad `` \quad x_2\,x_3 & `` \quad x_1 \quad `` \ ,
\end{array}
$$

Equation $(15)_1$ states that the vanishing of the invariant $A_{12} + A_{13}$ (equations 7) is the necessary and sufficient condition that the tangent plane, at x_1, to the ruled surface generated by the edge $x_1\,x_4$ shall meet the corresponding ruling of the opposite ruled surface in the point in which that ruling is met by the characteristic line of the plane u_1 $(y_1 = 0)$. We have described this line as the x_1 characteristic because it lies opposite the vertex x_1. The parenthesis indicates which subscripts appear in the geometrical statement and how often each one occurs.

The point in which $x_1\,x_4\,x_2'$ meets $x_2\,x_3$ is actually the point $(14)_2$ so that $(15)_1$ simply states under what condition the last named point coincides with the characteristic point $(10)_1$. We have preferred to state the relation as in $(15)_1$. By varying the subscripts similar relations can be written for the remaining edges of the tetrahedron.

If we remove from a tetrahedron a pair of opposite sides, the remaining edges constitute a unicursal four-line or space quadrilateral. Since we may remove any one of three different pairs of sides it follows that the tetrahedron may be thought of as giving rise to three distinct space quadrilaterals.

If we consider now the parentheses $(1^2\,2^2)$, $(2^2\,3^2)$, $(3^2\,4^2)$, and $(4^2\,1^2)$ we see that the edges they refer to form a space quadrilateral. It is to be noted also that the parentheses indicate a direction along this quadrilateral, namely, the direction pursued in going successively to the vertices x_1, x_2, x_3, and x_4.

The geometrical situations associated with each of the parentheses are given in statement $(15)_4$ and similar statements obtainable from the latter by rotating subscripts. It is readily verified that the accompanying invariant equalities are such that any three imply a fourth. We have accordingly,

THEOREM 3. *If any one of a one parameter family of space quadrilaterals be traced unicursally and if at three successive vertices the plane tangent at the vertex to the ruled surface generated by the ensuing edge meets the opposite edge in the point where the latter is met by the characteristic line corresponding to the successor vertex, then for the fourth vertex the property still persists.*

There is an important distinction between the theorem [illegible]
[illegible]

i. e., they required no hypothesis except that implied in the configuration. Theorem 3, on the other hand, stated that if a certain property recurred three times it would necessarily occur a fourth time. We can indicate this feature of the theorem by the fraction (3/4).

Most of the theorems of this paper will be like theorem 3. As to type they will be either (1/2), (2/3), (3/4), or (5/6). The possibility of each type existing is clear if we keep in mind the fact that there are four vertices, six edges, three edges in a plane, three edges on a vertex, while opposite edges occur in pairs.

The following considerations lead to a theorem of the (3/4) type. From any vertex an infinite number of tangent lines can be drawn to a ruled surface. We shall put our attention upon just one of these, namely, that one which meets the ruled surface in the ruling which corresponds to the vertex in question; *i. e.*, the ruling shall be given by that value of the parameter t which determines the vertex. From a given vertex three such lines can be drawn, one to each of the three ruled surfaces generated by the remaining three vertices taken in pairs. We shall speak of these as the "instantaneous" tangents from the vertex in question.

If we inquire under what conditions the "instantaneous" tangents from any vertex x_i are coplaner we find as necessary and sufficient condition

$$(16) \qquad M_i = N_i, \text{ (equations 7)}.$$

Of the four equalities (16) only three are independent. We have accordingly,

THEOREM 4. *If from three of the vertices of a variable tetrahedron the "instantaneous" tangents drawn to the ruled surfaces generated by the remaining edges are coplanar then for the fourth vertex the same property holds.*

9. *Developables.* The condition that the edge $x_i x_j$ shall generate a developable is simply the condition that the points x_i, x_j, x_i', and x_j' shall be coplaner. Applying this test to each edge in turn we find the following conditions for developables,

$$(17) \qquad \begin{array}{ll} (x_1 x_4)\ a_{12} a_{43} = a_{13} a_{42}, & (x_1 x_3)\ a_{32} a_{14} = a_{34} a_{12}, \\ (x_2 x_3)\ a_{21} a_{34} = a_{31} a_{24}, & (x_2 x_4)\ a_{41} a_{23} = a_{21} a_{43}, \\ (x_1 x_2)\ a_{13} a_{24} = a_{14} a_{23}, & (x_3 x_4)\ a_{42} a_{31} = a_{32} a_{41}. \end{array}$$

It is readily verified that only five of the equalities (17) are independent. We have, accordingly,

THEOREM 5. *If five of the edges of a variable tetrahedron generate developables the sixth edge also generates a developable.*

It is also readily verified that if the edges meeting in a vertex generate developables equation (16) holds. We have therefore

THEOREM 6. *If the three edges meeting in a vertex generate developables the instantaneous tangents from this vertex to the ruled surfaces generated by the opposite edges are coplanar.*

10. *Fixed Point and Plane.* If the coördinates of a point P referred to a fixed tetrahedron are P_i $(i=1, 2, 3, 4)$, the necessary and sufficient conditions that the point P shall be fixed are,

$$P_i' = \phi P_i, \ (i=1, 2, 3, 4). \tag{18}$$

Let us suppose that the coördinates of a point P referred to our variable tetrahedron as reference system, are a, b, c, and d. To find the conditions which make P a fixed point we must, in order to avail ourselves of (18), express the coördinates of P in terms of a fixed tetrahedron. For this purpose the obvious choice is to take as fixed tetrahedron the one in which the coördinates of the vertices x_1, x_2, x_3, x_4, of the variable tetrahedron are expressed. With this understanding

$$P \equiv a\,x_1 + b\,x_2 + c\,x_3 + d\,x_4. \tag{19}$$

If we substitute the expression (19) in (18) and treat the result as an identity in x_1, x_2, x_3, and x_4 we obtain the following conditions upon a, b, c, and d:

$$\begin{aligned} a' &= \phantom{-a_{12}a +{}} \phi a - a_{21}b - a_{31}c - a_{41}d, \\ b' &= -a_{12}a + \phi b - a_{32}c - a_{42}d, \\ c' &= -a_{13}a - a_{23}b + \phi c - a_{43}d, \\ d' &= -a_{14}a - a_{24}b - a_{34}c + \phi d. \end{aligned} \tag{20}$$

If we multiply the coördinates a, b, c, and d by the factor $\exp \int \phi dt$, we obtain equations (20) with the diagonal terms missing; that is, we obtain, except for a change in notation, the adjoint system (8). We can speak of either (20) or the equivalent canonical system as the "geometric adjoint"* of system (1) or the equivalent canonical system (5). In similar fashion we find that the coördinates of a fixed plane, after multiplication by a properly chosen factor, satisfy equations (5). We can summarize these results in

THEOREM 7. *The coördinates of a point which is fixed in space but referred to a variable tetrahedon, satisfy the system of equations which define*

* [illegible]

the face planes of the tetrahedron. The coördinates of a fixed plane, on the other hand, satisfy the defining system for the vertices of the tetrahedron.

The following considerations lead to an equivalent statement of Theorem 6. To this end it is necessary to recall the peculiar manner in which the defining system (1) was set up. The determinant Δ (equation 2) involves 16 functions arranged in a four-by-four array. In computing the coefficients in (1) we interpreted the elements in the i'th column of Δ as the coördinates of a point x_i. To state the same fact in a slightly dfferent form, suppose we are given the system of equations (1). A set of simultaneous solutions x_1, x_2, x_3 and x_4 is not interpreted as a point. On the contrary, if we have four such sets of solutions, the four values of x_i define the point x_i.

Looking back at equations (20) we see that in this instance it is precisely the simultaneous solutions which we interpret as the coördinates of a fixed plane referred to the variable tetrahedron. We have accordingly,

THEOREM 7_1. *If in a square array of 16 functions, with determinant different from zero, we interpret the elements in each column as the coördinates of a point, we obtain four points. As the parameter in which the functions are expressed, varies, the four points together with their joins generate a variable tetrahedron. If the elements in a given row are interpreted as the coördinates of a plane referred to the variable tetrahedron, then this plane is fixed in space. The four rows, accordingly, define four fixed planes.*

11. *Duo-flecnode Points.* With the variation of the parameter t, a pair of opposite edges of the tetrahedron will generate two ruled surfaces which are in one-to-one line correspondence. In connection with such a configuration, Professor Lane* has introduced the notion of straight line intersectors of the four skew lines made up of the corresponding rulings of the two surfaces together with their successors.

The two points in which a ruling of either surface is met by the intersectors in limiting position we shall denote as the "duo-flecnode points." The term flecnode is suggested by the similar situation in ordinary ruled surface theory. The prefix was added to distinguish from another case to be mentioned shortly.

Since any point on the edge $x_i x_j$ is given by the expression $x_i + \lambda x_j$, the determination of the flecnode points consists in determining appropriate values for λ. The following quadratics determine the values of λ for the edges indicated

* *Loc. cit.*, p. 291.

$$
\begin{aligned}
&(x_3\,x_4)-(a_{13}\,a_{41}+a_{23}\,a_{42})\lambda^2+(P+Q)\lambda+(a_{14}\,a_{31}+a_{32}\,a_{24})=0,\\
&(x_1\,x_2)-(a_{24}\,a_{41}+a_{31}\,a_{23})\lambda^2+(P-Q)\lambda+(a_{13}\,a_{32}+a_{14}\,a_{42})=0,\\
(21)\quad &(x_1\,x_4)-(a_{21}\,a_{42}+a_{21}\,a_{43})\lambda^2+(R+P)\lambda+(a_{24}\,a_{12}+a_{13}\,a_{34})=0,\\
&(x_2\,x_3)-(a_{12}\,a_{31}+a_{34}\,a_{42})\lambda^2+(R-P)\lambda+(a_{13}\,a_{21}+a_{24}\,a_{43})=0,\\
&(x_1\,x_3)-(a_{21}\,a_{32}+a_{34}\,a_{41})\lambda^2+(R-Q)\lambda+(a_{23}\,a_{12}+a_{14}\,a_{43})=0,\\
&(x_2\,x_4)-(a_{12}\,a_{41}+a_{32}\,a_{43})\lambda^2+(R+Q)\lambda+(a_{14}\,a_{21}+a_{23}\,a_{34})=0,
\end{aligned}
$$

where

$$P=A_{24}-A_{13},\quad Q=A_{14}-A_{23},\quad R=A_{34}-A_{12}.$$

By observing the form of the coefficient of λ in equations (21) we see that the following theorems are true:

THEOREM 8. *If on two of the edges of a variable tetrahedron, the duo-flecnode points are harmonic with respect to adjacent vertices it follows that the property holds for the third edge coplanar with the first two.*

THEOREM 9. *If on one, and only one, line of each of two pairs of opposite edges, the flecnode points are harmonic with respect to adjacent vertices it follows that on one, and only one, line of the remaining pair of edges are the flecnode points harmonic with respect to adjacent vertices.*

12. *Tri-flecnode Points.* If we have two ruled surfaces with generators in one-to-one correspondence we can obtain another set of points similar to the ones we have designated as duo-flecnode points.* This time we make a selection of four skew lines by choosing on the first surface a ruling together with two successor rulings. The corresponding ruling on the second surface constitutes a fourth line. The two straight line intersectors of the four lines take a limiting position as the successor rulings on the first surface approach coincidence with the ruling itself.

The points where the ruling of the first surface is met by the intersectors in limiting position we shall designate as the tri-flecnode points on that ruling. On the other hand, the points where the ruling of the second surface is met by these same intersectors will be designated as the uno-flecnode points on that ruling. In the equations below, the quadratic in λ determines the tri-flecnode points on $(x_1\,x_4)$, while the similar quadratic in μ determines the corresponding uno-flecnode points on ([illegible]).

$$
\begin{aligned}
(22)\quad &\lambda^2(a_{12}C_4-a_{13}B_4)+\lambda(a_{12}C_1-a_{13}B_1-a_{12}C_4+a_{13}B_4)+(a_{12}C_1-a_{13}B_1)=0,\\
&\mu^2(a_{12}B_4-a_{42}B_1)+\mu(a_{42}C_1-a_{12}B_4-a_{12}C_4+a_{13}B_1)+(a_{13}C_4-a_{13}C_1)=0.
\end{aligned}
$$

The expressions A_i, B_i, C_i, D_i are defined by the equation,

$$x_i'' = A_i x_1 + B_i x_2 + C_i x_3 + D_i x_4.$$

By permuting subscripts similar quadratics can be written for the remaining edges.

13. *Tri-flecnode Points Indeterminate.* For the tri-flecnode points on $(x_1 x_4)$ to be indeterminate it is necessary and sufficient that the coefficients of λ^2, λ^1, and λ^0 in $(22)_1$ shall be zero. We must have,

$$(23) \qquad \begin{gathered} a_{42} C_4 = a_{43} B_4, \;\; a_{12} C_1 = a_{13} B_1, \\ a_{42} C_1 + a_{12} C_4 - a_{43} B_1 + a_{13} B_4. \end{gathered}$$

The third of equations (23) when taken in conjunction with the other two leads to two distinct possibilities. The first term in the left member may be equated to either the first or second term in the right member. We have, therefore, as a consequence of (23), either the following set of conditions

$$(24) \quad a_{42} C_4 = a_{43} B_4, \; a_{12} C_1 = a_{13} B_1, \; a_{42} C_1 = a_{13} B_4, \; a_{12} C_4 = a_{43} B_1,$$

or

$$(25) \quad a_{42} C_4 = a_{43} B_4, \; a_{12} C_1 = a_{13} B_1, \; a_{12} C_4 = a_{13} B_4, \; a_{42} C_1 = a_{43} B_1.$$

From equations (24) we readily find that the coefficients of μ^2, μ^1, μ^0 in $(22)_2$ are zero. In other words, if equations (24) hold, the edge $x_2 x_3$ lies on the osculating quadric to the ruled surface generated by the opposite edge $x_1 x_4$.

On the other hand, equations (25) are the necessary and sufficient conditions that the edge $x_1 x_4$ should generate a plane. In fact, if we insist that the osculating plane to the curve traced by x_1 shall be identical with the corresponding one for the curve traced by x_4, we obtain the last two of equations (25) together with a third equation which can be obtained from the first two of equations (25).

If, then, equations (25) hold the edge x_1 x_4 generates a plane. The plane thus determined will intersect the opposite edge in a point. All the lines lying in the plane will meet the edge $x_1 x_4$ and its two successor rulings as well as the edge $x_2 x_3$. The indeterminateness of the tri-flecnode points is obvious in this case.

If equations (24) hold, the points determined by $(22)_2$ are also indeterminate so that the ruling $x_1 x_4$, together with its two successor rulings and the opposite ruling $x_2 x_3$ are all in quadric position. We can summarize this section as,

THEOREM 9. *If two ruled surfaces are in one-to-one line correspondence, indeterminateness of the tri-flecnode curves on one of the two surfaces*

*implies one of two things, the surface on which the curves are indeterminate is a plane, or else the second surface belongs to the flecnode congruence of the first.**

From equations (25) we can readily derive equation $(17)_1$, which is the condition that $(x_1 x_4)$ should generate a developable. By rotating letters and subscripts we obtain from (25) the conditions that each edge in turn shall generate a plane. In each case only three of the four equations are independent. We shall choose as independent conditions the first two in each set together with the condition for developability. We already know that the latter conditions are so related that any five imply the sixth. It also turns out that the first two in each set are so related that if they hold for five edges they necessarily hold for the sixth edge. We have then,

THEOREM 10. *If five of the edges of a variable tetrahedron generate planes, the sixth edge also generates a plane.*

CONCLUSION.

We have considered in this paper a one parameter family of tetrahedrons. The various parts of the configuration have been seen to be interdependent in such a fashion that properties of a certain portion of the configuration implied like properties for the rest of the configuration.

Only a few of the simplest properties have been considered.

If we wish to apply the results of this paper to the variable reference tetrahedron as it occurs in various branches of Projective Differential Geometry we need merely to have for the vertices of the tetrahedron such a defining system as (1) or (5).

If we turn to the theory of space curves, we find that equations (5), p. 239 of Wilczynski's *Proj. Diff. Geom.*, are just such a system as we require. In fact, if, as is usually the case, p_1 is set equal to zero, equations (5) are in what we have designated as the canonical form.

In general surface theory the points y, y_u, y_v, and y_{uv} frequently constitute the reference tetrahedron. If the parameter u alone varies a one parameter family of tetrahedrons is obtained. There is no difficulty in writing down a canonical system for these points.

A succeeding paper will extend these results to four-space.

University of Washington, Seattle, Wash.

On the Moduli of Algebraic Functions Possessing a Given Monodromie Group.

By Oscar Zariski.

1. It is known that, given a transitive group G of degree n generated by a set of substitutions $S_1, S_2, \cdots, S_\omega$, such that $S_1 S_2 \cdots S_\omega = 1$, there exist infinite algebraic functions y of x of order n possessing G as monodromie group and having ω branch points $\alpha_1, \alpha_2, \cdots, \alpha_\omega$, which can be assigned arbitrarily, and such that for a given system of loops in the x plane the substitution on the branches of y at α_i is exactly S_i. By letting the position of the branch points α_i vary arbitrarily, we obtain a *complete continuous system,* which we will call Σ, of algebraic functions, which depend on $\omega - 3$ essential parameters—the cross-ratios of the set (α). All functions of Σ are of the same order n and have the same monodromie group G and the same genus p. The value of p depends only on the assigned substitutions S_i, viz., if S_i is composed of μ_i cycles of orders $\nu_{1i}, \nu_{2i}, \cdots, \nu_{\mu_i i}$ respectively, then

$$2n + 2p - 2 = \sum_{i=1}^{\omega} \left(\sum_{j=1}^{\mu_i} \nu_{ji} - \mu_i \right). \tag{1}$$

We now ask: *Under what conditions will the functions of Σ be of general moduli?* By saying that an n-valued algebraic function y is of general moduli, we mean that the algebraic curve ϕ, defined in the invariantive sense of Algebraic Geometry, and on which y corresponds to a definite linear involution $g_n{}^1$, is of general moduli. Evidently a *necessary* condition, that the functions y of Σ be of general moduli, is that the above number $\omega - 3$ of essential parameters should not be less than the number $3p - 3$ of the birational moduli of the general curve of genus p. Hence, necessarily

$$\omega \geqq 3p. \tag{2}$$

If (2) *holds, will it be possible to affirm that the generic function y of Σ is of general moduli?* There is at least one case in which this is not necessarily true, namely when G is an imprimitive group. In fact, in this case, any function y of Σ defines on the corresponding curve ϕ a $g_n{}^1$ *composed of an involution.* If this involution is irrational and $p > 1$, then it is well known that ϕ, or y, is of special moduli. The following example may serve as an illustration: Let $\omega = 6$, and let

$$S_1 = S_2 = S_3 = S_4 = (13)\ (24)\,;\ S_5 = S_6 = (12). \tag{3}$$

The functions y of Σ are of genus 2 and the number of essential parameters on which they depend is 3, the same as that of the birational moduli in the case $p = 2$. However, we prove, that the functions y are of special moduli. In fact, the group G of degree 4 and of order 8 generated by the substitutions (3), is imprimitive, the sets of imprimitivity being [1, 2]; [3, 4]. Hence the g_4^1 is composed of an involution γ_2^1. If ψ is an algebraic curve in one-to-one correspondence with the sets of the γ_2^1 on ϕ, then to the g_4^1 on ϕ there corresponds on ψ a g_2^1, and hence a two-valued algebraic function z, whose branch points are evidently among the branch points of y. Now, the substitutions S_1, S_2, S_3, S_4 permute the sets of imprimitivity, while S_5 and S_6 leave invariant each set. Hence z, a 2-valued algebraic function with 4 branch points α_1, α_2, α_3, α_4, is of genus 1. It follows, that the γ_2^1 on ϕ is an elliptic involution, and that therefore the functions y of Σ are of special moduli.

Excluding the obvious case of imprimitive groups, the above formulated question constitutes one of the most general problems, which arise in connection with the theory of Galois as applied to Algebraic Geometry. The affirmative answer, i. e. the theorem that if G is a primitive group and if (2) holds, then the functions of Σ are of general moduli, is probably correct, although at present we are not able to prove it rigorously. The theorem which we establish in this paper (section 3), and which generalizes a theorem stated by Severi for the case in which the assigned substitutions S_i are transpositions,* may be considered as the first step toward the solution of the problem. It gives an important geometric criterion, which reduces this group problem to a problem of the theory of virtually complete series on an algebraic curve.

2. Our problem may be considered from a slightly different point of view, which we are going to explain. On a given curve ϕ of genus p and of general moduli there exist, for a sufficiently large given value of n, an infinite class of linear involutions g_n^1. The monodromie group of *a generic* g_n^1 of the class, i. e. the monodromie group of the corresponding n-valued algebraic function, is the total group. But for special g_n^1's of the class the group may well reduce to a subgroup of the total group. This will generally happen when and only when each multiple point of the g_n^1 is of order > 2, so that the Jacobian set of the g_n^1 does not contain simple points. We call a transitive group G of degree n *non-special of genus p*, if G is the monodromie group of *some* g_n^1 on a curve ϕ of genus p *and of general moduli*. Otherwise we say that G *is special of genus p*. All transitive groups of a given degree are

* F. Severi, "Sul Teorema [illegible]

thus divided, *with respect to a given value p,* into two classes, special and non-special groups of genus p. It follows from the definition that, if a group G of degree n is special of genus p, then no matter how we specialize the $g_n{}^1$ on a curve of genus p of general moduli, by letting its Jacobian set acquire multiple points, we never obtain a $g_n{}^1$ possessing G as monodromie group. Our problem amounts to finding a criterion which should enable us to decide whether a given transitive primitive group G is special or non-special of genus p.

Let, as in section 1, a set of substitutions $S_1, S_2, \cdots, S_\omega$ be introduced, which generate G and such that $S_1 S_2 \cdots S_\omega = 1$. We say then, that we have *a representation of genus p and of rank ω of the group G,* where p is defined by (1). A group G is certainly special of genus p, if in *any representation* of genus p of G, the rank ω is less than $3p$. The most important example of such groups is given by the primitive solvable groups. The author showed elsewhere* that if $p > 6$, then in any representation of genus p of a primitive solvable group the rank ω is less than $3p$, and thus deduced the conclusion that an algebraic equation of genus $p > 6$ and of general moduli cannot be solved parametrically in radicals.

On the other hand, if a transitive primitive group G does admit a representation of genus p and of rank $\omega \geqq 3p$, it does not yet follow, although it seems probable (see section 1), that G is non-special of genus p. If this were true, we would have the following remarkable fact: that the special groups of a given genus can be completely characterized by an arithmetical condition.

3. If the assigned substitutions S_i are transpositions, so that G is the total group of degree n, it is well known that G is special of genus p, if and only if $n < p/2 + 1$. A theorem due to Severi (see quoted paper) says that the most general curve of genus p containing a $g_n{}^1$ depends on $3p - 3 - i$ moduli, where i is the *index of speciality* of the complete two-fold series $|2g_n|$. The purpose of this paper is to show that the method employed by Severi can still be successfully employed in the case in which the assigned substitutions are (not transpositions, but) arbitrary. We thus obtain a theorem similar to that of Severi, in which, however, the ordinary complete series $|2g_n|$ is replaced by a conveniently defined *virtually complete series* $||2g_n||$. In order to be able to state our result, we anticipate some of the definitions and statements of the later sections.

---

* Sull'impossibilità di risolvere parametricamente per radicali un'equazione algebrica $f(x, y) = 0$ di genere $p > 6$ a moduli generali, *Rendiconti della R. Accademia dei Lincei* (6), Vol. 3 (1926).

Let y be a generic function of Σ, and let $g_n{}^1$ be the linear involution on the corresponding algebraic curve ϕ. Let $\Gamma_1, \Gamma_2, \cdots, \Gamma_\omega$ be the sets of the $g_n{}^1$ which contain multiple points of the involution. Then (see section 1) Γ_i contains one multiple point P_{1i} of order ν_{1i}, one multiple point P_{2i} of order $\nu_{2i}, \cdots$, one multiple point $P_{\mu_i i}$ of order $\nu_{\mu_i i}$ and $n - \nu_{1i} - \nu_{2i} - \cdots \nu_{\mu_i i}$ simple points. Consider the ω sets

$$(4) \qquad (P_{1i}{}^{\nu_{1i}-1}, \quad P_{2i}{}^{\nu_{2i}-1}, \cdots, \quad P_{\mu_i i}{}^{\nu_{\mu_i i}-1}).$$

Given a linear series $g_m{}^r$ on ϕ, the above sets are said to be *neutral sets* of the series, if each set imposes one condition only on the sets of the $g_m{}^r$ constrained to contain it, i. e. if the sets of the $g_m{}^r$ which contain one point of one of the sets (4), say P_{1i}, consequently contain each point P_{ji} of this set to the multiplicity $\nu_{ji} - 1$.

If the $g_m{}^r$ is not contained in a larger series of the same order, possessing the same ω neutral sets, then the series is said to be virtually complete.

A set A of m points of which none coincides with any of the points P_{ji}, is contained in a unique virtually complete series $\| A \| = \| g_m \|$.

For reasons which will be explained later, the *index of speciality* i of a virtually complete series $g_m{}^r$ is defined as follows:

$$i = r - [m - p - \sum_{i=1}^{\omega} (\sum_{j=1}^{\mu_i} \nu_{ji} - \mu_i)] = r - [\omega - 3p + 2 + m - 2n].$$

A virtually complete series is *special* or *non-special* according as i is > 0 or $= 0$.

In particular, a virtually complete series of order $2n$ is a series $g_{2n}{}^{\omega-3p+2+i}$, where i is the index of speciality of the series.

We can now announce our result as follows:

Let Γ and Γ' be two arbitrary sets of the $g_n{}^1$, which corresponds to a generic function y of Σ, and let $\| g_{2n} \| = \| \Gamma + \Gamma' \|$ be the virtually complete series defined by the set $\Gamma + \Gamma'$. The functions of Σ depend on $3p - 3 - i$ moduli, where i is the index of speciality of the series $\| g_{2n} \|$.

In particular, we have:

The necessary and sufficient condition that the functions of Σ [illegible] moduli (and hence, that the group G be non-special of genus p), is that the above virtually complete series $\| g_{2n} \|$ should be non special.

g_m^1 on ϕ, we transform ϕ into a plane algebraic curve C of order $m+n$ with two fixed multiple points A and B of orders m and n respectively. The curve C possesses other variable multiple points equivalent to $(m-1)(n-1)-p$ ordinary double points. The linear series g_n^1 and g_m^1 are cut out on C by the lines of the pencils A and B respectively. By letting the ω branch points α of the function y and the g_m^1 on ϕ vary, we obtain a continuous system of curves C, which we will denote by K. *We say that if m is taken sufficiently large, the system K is irreducible, and that the generic curve C of K possesses ordinary double points only, and that moreover none of the ω tangents to C through A passes through a double point of C.*

To see this, let us take $m \geqq 2p+2$, and let us consider a generic complete series g_m^{m-p} on ϕ. We imagine the sets of the series represented by the points of a projective space S_{m-p}, in which the lines are the images of g_m^1's contained in the g_m^{m-p}. As an immediate consequence of the theorem of Riemann-Roch it follows, that the g_m^{m-p} has no fixed points, and that its residual series with respect to any 3 (distinct or coincident) points, is a g_{m-3}^{m-p-3}.

We consider in S_{m-p} the following loci:

1. The locus V_1 of points, images of the sets of the series g_m^{m-p} which contain a variable triple of points of ϕ, belonging to a set of the g_n^1.

2. The locus V_2 of points, images of the sets of the g_m^{m-p}, which contain a pair of (distinct or coincident) points belonging to one of the ω sets of the g_n^1, which contain multiple points of this last series.

The loci V_1 and V_2 consist of a finite number of varieties, and by the foregoing remark each of these varieties is of dimensions $\leq m-p-2$. Hence, it is sufficient to consider a generic line in S_{m-p} which does not meet the loci V_1 and V_2, in order to have a g_m^1 on ϕ which together with the g_n^1 gives rise to a curve C of K, satisfying the required conditions.

That the system K is irreducible, follows immediately from the fact that both the system Σ and the variety of all g_m^1's on a curve of genus p are irreducible.

To evaluate the *dimension* of K, we observe that, when the ω tangents to C through A are assigned, the projective correspondence between the sets of the g_n^1 and the lines of the pencil A is uniquely determined,* while there are still ∞^3 ways of setting up a projective correspondence between the sets of the g_m^1 and the lines of the pencil B. Since the dimension of the variety of the linear involutions g_m^1 on a curve of genus p is $2m-p-2$, it follows that K is of dimension

* We suppose that $\omega \geqq 3$, since, if $\omega = 2$, the curves C are rational.

(5) $$r = 2m - p - 2 + 3 + \omega = 2m - p + \omega + 1.$$

The peculiarity of the system K consists in the fact that each of the ω tangents through A to a generic curve C of K is, in general, a multiple tangent, having with C one or more contacts of order $\geqq 1$. From the foregoing considerations it follows that the points of contact of any of the ω tangents are *distinct simple points* of C. Thus, if we denote the tangent corresponding to the branch point α_i by the same letter α_i, we have that α_i has with C one contact of order ν_{1i} at a simple point P_{1i}, one contact of order ν_{2i} at a simple point $P_{2i}, \cdots$, one contact of order $\nu_{\mu_i i}$ at a simple point $P_{\mu_i i}$.

5. Following the method of Severi, we consider the $\sigma = 2m + 2p - 2$ tangents through B to a generic curve C. Let us denote them by λ. If of the σ tangents λ, $\sigma - i$ $(i \geqq 0)$ only can be assigned arbitrarily, then the curves of K depend on $3p - 3 - i$ moduli. Let C_0 be a generic but fixed curve of K and let us consider the variety of curves of K which touch the tangents λ. Let H be that *irreducible* part of this variety which contains C_0. Since, by hypothesis, the sets λ depend on $\sigma - i$ arbitrary parameters, and since the dimension of K is $2m - p + \omega + 1$, it follows that the *dimension* of H is

(6) $$r' = (2m - p + \omega + 1) - (2m + 2p - 2 - i) = \omega - 3p + 3 + i,$$

and that the characteristic series of H on C_0 is of dimension

(7). $$r' - 1 = \omega - 3p + 2 + i.$$

In order to characterize the above series, we prove the following:

Lemma. *If in an irreducible continuous system $\{D\}$ of plane algebraic curves, the generic curve D has μ contacts of orders $\nu_1, \nu_2, \cdots, \nu_\mu$ at simple points $P_1, P_2, \cdots, P_\mu$ with a variable line α passing through a fixed point A, then the characteristic series of the system on D possesses $[P_1^{\nu_1 - 1}, P_2^{\nu_2 - 1}, \cdots, P_\mu^{\nu_\mu - 1}]$ as a neutral set of points.*

Let D_0 be a generic fixed curve of $\{D\}$. To simplify the proof, we may suppose that the point A is at infinity on the y-axis, and that the multiple tangent α to D_0 coincides with the y-axis. We consider a generic algebraic system ∞^1 of curves D, containing D_0. The equation of a generic curve D of the system is of the form

[illegible]

where ξ and b_i are functions of t. We suppose that D_0 corresponds to $t=0$, so that the equation of D_0 reduces to

$$f(x,y,0)=x\phi_0(x,y)+(y+b_1{}^0)^{\nu_1}(y+b_2{}^0)^{\nu_2}\cdots(y+b_\mu{}^0)^{\nu_\mu}\psi_0(y),$$

where

$$\phi_0(x,y)=\phi(x,y,0),\quad b_i{}^0=\lim_{t=0}\cdot b_i,\quad \psi_0(y)=\psi(y,0),$$

and

$$\phi_0(0,-b_i{}^0)\neq 0,\qquad \psi_0(-b_i{}^0)\neq 0\qquad (i=1,2,\cdots,\omega).$$

The characteristic group on D_0 is cut out by the curve

$$\begin{aligned}(8)\qquad [\partial f(x,y,t)/\partial t]_{t=0} &= [\phi_0(x,y)+x(\partial\phi_0/\partial x)]\ (d\xi/dt)_0\\ &+x(\partial\phi/\partial t)_0+(y+b_1{}^0)^{\nu_1-1}(y+b_2{}^0)^{\nu_2-1}\cdots(y+b_\mu{}^0)^{\nu_\mu-1}\bar\psi(y)=0,\end{aligned}$$

where

$$\begin{aligned}(9)\qquad \bar\psi(y) &= (y+b_1{}^0)(y+b_2{}^0)\cdots(y+b_\mu{}^0)\\ &\times\{\psi(y,0)\,[\nu_1/(y+b_1{}^0)+\nu_2/(y+b_2{}^0)+\cdots+\nu_\mu/(y+b_\mu{}^0)]\\ &+(\partial\psi/\partial t)_0\},\end{aligned}$$

and where $(\partial\phi/\partial t)=[\partial\phi(x+\xi,y,t)/\partial t]$ is evaluated by keeping ξ constant. Suppose now that the characteristic group contains the point $x=0$, $y=-b_1{}^0$. Then, since $\phi_0(0,-b_1{}^0)\neq 0$, it follows from (8) and (9) that $(d\xi/dt)_0=0$, and hence that the curve (8) passes through each point $x=0$, $y=-b_i{}^0$ and has at that point a (ν_i-1)-fold contact with the y-axis, q. e. d.

By this lemma it follows that the characteristic series of H on C_0 possesses ω neutral sets:

$$(4')\qquad [P_{1i}{}^{\nu_{1i}-1},\ P_{2i}{}^{\nu_{2i}-1},\cdots,\ P_{\mu_i i}{}^{\nu_{\mu_i i}-1}],\qquad (i=1,2,\cdots,\omega).$$

Moreover, the curves of H, infinitely near to C_0, pass through the double points of C_0 and through the σ points of contact of the tangents λ with C_0. Hence, as in the case considered by Severi, the above characteristic series is of order $2n$ and is totally contained in the two-fold of the series $g_n{}^1$, cut out by the lines of the pencil A.

The characteristic series of H on C_0 can be thought of as *a virtually complete series of index of speciality i.* To show this, we must give a general definition of a virtually complete series, and we must also explain what is meant by a special virtually complete series and by the index of speciality of such a series.* This is done in the next section.

* The general theory of virtual series was first developed by Nöther in "Über die Schnittpunktsysteme einer algebraischen Curve mit nicht-adjungirten Curven," *Mathematische Annalen*, Vol. 15 (1879), p. 507-528. In this paper Nöther uses as

6. A linear series $g_m{}^r$ on a curve ϕ, with ω neutral sets

$$(10) \qquad [P_{1i}{}^{k_{1i}}, \quad P_{2i}{}^{k_{2i}}, \cdots, \quad P_{\mu_i i}{}^{k_{\mu_i i}}], \qquad (i = 1, 2, \cdots, \omega)$$

is *virtually complete,* if it is not contained in a series of the same order and greater dimension, possessing the same ω sets (10) as neutral sets. A virtually complete series will be denoted by $\| g_m \|$.

A set G_m of m points of ϕ, none of which coincides with a point P_{ji} of the above sets, belongs to a unique virtually complete series. Due to this theorem, this series can be denoted by $\| G_m \|$.

To prove this, let $g_m{}^r$ $(r \geqq m - p)$ be the complete series to which the set G_m belongs. In the representative space S_r, the points of which correspond to the ∞^r sets of the $g_m{}^r$, let the point O correspond to the set G_m, and let S_{r_i} be the linear space in S_r the points of which correspond to the sets of the $g_m{}^r$, which contain the points $P_{1i}, P_{2i}, \cdots, P_{\mu_i i}$ to the multiplicities $k_{1i}, k_{2i}, \cdots, k_{\mu_i i}$ respectively. By hypothesis, the spaces S_{r_i} $(i = 1, 2, \cdots, \omega)$ do not contain the point O. Any linear series $g_m{}^\sigma$, containing G_m, and having the above ω neutral sets of points, is represented in S_r by a linear space, which evidently belongs to each of the spaces S_{r_i+1} determined by O and S_{r_i}. Hence any such series is contained in a unique series of the greatest dimension, which is represented in S_r by the space S_ρ — intersection of the ω spaces S_{r_i+1}. *This series $g_m{}^\rho$, for which the above ω sets* (10) *are neutral sets of points, is the virtually complete series, uniquely determined by the set G_m.*

The virtually complete series $g_m{}^\rho$, or $\| G_m \|$, is said to be non-special, if the following conditions are satisfied:

---

a starting point for his theory the definition of non-adjoint, or virtually adjoint curves, and derives two distinct types of virtually complete series on a given curve f. Series of the first type are cut out on f by complete linear systems of curves, subject only to the condition of being virtually adjoint curves of f. Series of the second type are cut out on f by complete linear systems of virtually adjoint curves, which are "*gleichsingulär,*" this being the term which Nöther uses in order to express the fact that the curves of the system in addition satisfy certain contact conditions at each multiple point of f. It clearly appears, although it is not pointed out explicitly by Nöther, that these virtual series of the second type possess neutral sets of points. [illegible] this property. The procedure [illegible] rapidly to the fundamental theorems of the theory, avoiding the complicated analysis relative to the behaviour of "*gleichsingulär*" virtually-adjoint curves at the singular points of the curve f.

For a treatment of virtual series with neutral *pairs*, see F. Severi, *Vorlesungen*

1.° *The ordinary complete series $g_m{}^r$, or $| G_m |$, is non-special, and hence* $r = m - p$.

2.° *The points P_{ji} of the i-th neutral set, taken to the corresponding multiplicities k_{ji}, impose $k_{1i} + k_{2i} + \cdots + k_{\mu_i i}$ independent conditions on the sets of the series $g_m{}^r$, constrained to contain them;* hence,

$$r_i = m - p - k_{1i} - k_{2i} - \cdots - k_{\mu_i i} \geqq 0. \tag{11}$$

3.° *The ω spaces S_{r_i+1} are linearly independent and so intersect in a space of dimension*

$$\rho = (r_1 + 1) + (r_2 + 1) + \cdots + (r_\omega + 1) - (\omega - 1)r. \tag{12}$$

Putting

$$\sum_{i=1}^{\omega} \sum_{j=1}^{\mu_i} k_{ji} = k,$$

and taking into account (11), (12) becomes

$$\rho = m - p - k + \omega. \tag{13}$$

Conversely, if the dimension ρ of a virtually complete series $\| G_m \|$ is given by (13), the conditions 1°, 2° and 3° all hold, and hence the series is non-special.

If one, at least, of the above conditions is not satisfied, the virtually complete series $\| G_m \|$ is *special.* The dimension of a special virtually complete series $g_m{}^\rho$ is given by the formula

$$\rho = m - p - k + \omega + i, \tag{14}$$

where i is > 0. *The number i is called the index of speciality of the series.**

* The above definitions suggest that it is possible, as in the case of neutral pairs, to consider a virtually complete series as the limit of an ordinary complete series on a variable curve of genus $P = p + k - \omega$. It will be noticed however, that if $k > 2\omega$, i. e. if not all the neutral sets are neutral pairs, the upper limit $2p - 2 + k$ [See (15)] of the order of a special series $\| g_m \|$ is less than $2P - 2$, and that hence a special virtually complete series of the highest order is *not* the limit of the canonical series g_{2P-2} on a variable curve of genus P. For instance, if we consider series having one neutral triple ($k = 3$), the special series of the highest order are series g_{2p+1}, and these are limits of series contained in the canonical series g_{2p+2}^{p+1} on the variable curve of genus $p + 2$. If a curve of genus p is represented by a plane curve f_0 of order m with d ordinary double points and one ordinary triple point O, any of the above series g_{2p+1} is cut out by a system of (*virtually adjoint*) curves of order $m - 3$, defined by the condition of passing *simply* through the double points of f and through the point O and of having at O a fixed tangent. If f is a variable curve of genus $p + 2$ with $d + 1$ double points, which tends to f_0, the curves of the above system are limits of the adjoint curves of f, passing through a fixed point O', which, as f tends to f_0, approaches O on a definite direction.

It is easily shown that the order m of a special virtually complete series g_m^ρ cannot surpass $2p - 2 + k$. In fact, the residual series of the g_m^ρ with respect to the neutral sets is of order $m - k$ and of dimension $\rho - \omega = m - k - p + i$. Hence the residual series is special in the ordinary sense, and therefore

$$(15) \qquad m \leqq 2p - 2 + k.$$

Without going into the details of the proof, we observe that *there exist special virtually complete series of the highest order $2p - 2 + k$.* But, while in the case of neutral pairs of (distinct or coincident) points, there exists only one such series, viz., *the virtual canonical series,* in any other case this is not true any more. If $k > 2\omega$, there exist infinite special virtual series of the highest order, each of which may be considered as *a* virtual canonical series. However, the following property still holds: *a special virtual series is always contained in one of the virtual canonical series.*

7. The characteristic series of the system K on C_0 is a $g^{2m-p+\omega}_{2n+2m+2p-2}$ (see sec. 4), possessing the ω neutral sets (4) of sec. 3. We have in this case

$$k = \sum_{i=1}^{\omega} \sum_{j=1}^{\mu_i} (\nu_{ji} - 1) = 2n + 2p - 2,$$

and hence the highest order of a special virtual series is

$$k + 2p - 2 = 2n + 4p - 4.$$

Since, by hypothesis, $m > 2p + 2$, the above characteristic series is *non-special.* Using the formula (13), which gives the dimension of a virtually complete non-special series, we find

$$\rho = (2m + 2n + 2p - 2) - p - k - \omega = 2m - p + \omega.$$

It follows, that the characteristic series of K on C_0 is virtually complete.

With regard to this result, we want to make the following remark. The noted theorem, announced by F. Enriques in 1904,* that the characteristic series of a complete continuous system of algebraic curves is complete, is absolutely true if the curves of the system have ordinary double points only. However, it was pointed out by F. Severi † that if the curves of the system possess higher singularities, the characteristic series may not be complete.

* This is a particular case of a more general theorem, due to Enriques, "Sulla proprietà caratteristica delle superficie irregolari," *R. Accademia delle Scienze di Bologna*, Vol. 1 (1904), p. 5.

For instance, in the case recently considered by B. Serge* of a continuous system of curves possessing a variable tacnode, the characteristic series is not complete. However, the above result regarding the characteristic series of the system K suggests the possibility of restoring, at least formally, the validity of Enriques' theorem, by considering the *characteristic series* as a *virtual* series. For instance, in the above example considered by Segre, the characteristic series on a generic curve of the system is easily seen to possess the two points on the branches of the tacnode as a neutral pair, and as such the characteristic series is *virtually complete.* In the quoted paper Segre also considers a system of curves having a variable ordinary r-fold point P. Here, if we denote by $P_1, P_2, \cdots, P_n$ the n points of the curve on the branches passing through P, it is easily seen that *the residual series* of the characteristic series with respect to any of the points P_i possesses the remaining $n-1$ points as a neutral set, and as such this residual series is virtually complete. We do not intend here to pursue the investigation of this interesting question. We are satisfied with having called attention to an interpretation of Enriques' theorem, which may be useful for the study of continuous systems of curves.

8. We now come back to the characteristic series of the system H on C_0. We know already (see section 5) that the order and the dimension of this series are $2n$ and $\omega - 3p + i + 2$, and that the sets (4) of sec. 3 are neutral sets of the series. If i' is the index of speciality of the characteristic series, then the dimension of the virtually complete series $\| g_{2n} \|$ is [see (14), sec. 6]

$$2n - p - k + \omega + i' = \omega - 3p + 2 + i'.$$

Hence $i' \geqq i$, *where the sign* $=$ *holds, if the above characteristic series is virtually complete.* Now, it can be proved that in fact i' *is* equal to i, by repeating almost literally the reasoning of Severi (see quoted paper in sec. 3). It is only necessary to observe that the theorem that the residual series of a complete series with respect to a fixed set of points is complete, *holds also in the case of virtually complete series.* It follows, as in the case considered by Severi, that the σ points of contact of the tangents λ to C_0 through B impose $\sigma - i'$ conditions only to the curves of K infinitely near to C_0. Therefore etc.

To complete the proof of the theorem announced in sec. 3, we have only to show that the characteristic series of the system H in C_0 contains each set $\Gamma_1 + \Gamma_2$ made up of two sets of the g_n^1 cut out on C_0 by the lines of the pencil A. To show this, let A and B be taken at infinity on the x and y axes

* "Sui sistemi continui di curve piane con tacnodo, *Rendiconti della R. Accademia dei Lincei* (6), Vol. 9 (1929).

respectively, and let $f(x, y) = 0$ be the equation of C_0. Among the curves of K infinitely near to C_0 we find evidently the transforms of C_0

$$f[(x + a(t), \quad y + b(t)] = 0,$$

by an infinitesimal translation. The corresponding characteristic sets are cut out by the polar curves of C_0,

$$a_1(\partial f/\partial x) + b_1(\partial f/\partial y) = 0,$$

with respect to the points on the line at infinity. In particular, the characteristic set cut out on C_0 by the polar curve $\partial f/\partial y = 0$ is made up of the $\sigma = 2m + 2p - 2$ points of contact of the tangents λ and of the n points $A_1, A_2, \cdots, A_n$, *each counted twice,* which lie on the n linear branches of C_0 through A. The points $A_1, A_2, \cdots, A_n$ form evidently a set of the $g_n{}^1$. Hence, the characteristic series $\| g_{2n} \|$ contains the set $[A_1{}^2, A_2{}^2, \cdots, A_n{}^2]$. Since the ∞^2 sets of $2n$ points made up of 2 arbitrary sets of the $g_n{}^1$ form a linear series $g_{2n}{}^2$ possessing the same ω neutral sets as the virtually complete series $\| g_{2n} \|$, and since both series contain a common set $[A_1{}^2, A_2{}^2, \cdots, A_n{}^2]$, we deduce that the above $g_{2n}{}^2$ is contained in the series $\| g_{2n} \|$.

9. In order to apply the theorem of sec. 3 to the group problem announced in sec. 1, it would now be necessary to investigate the virtually complete series $\| g_{2n} \|$, with special reference to the question of its index of speciality. In the case $\omega \geqq 3p$ (assuming always that G is a primitive group), the aim of this investigation would be to prove, if possible, that the above series is non-special in general, i. e. when the ω branch points are arbitrary. In the case $\omega < 3p$, the above series is certainly special, since it contains the $g_{2n}{}^2$ obtained by combining the sets of the $g_n{}^1$ two at a time. Hence, if i is the index of speciality of the complete series $\| g_{2n} \|$, then

$$\omega - 3p + 2 + i \geqq 2,$$

or

$$i \geqq 3p - \omega.$$

If the sign $=$ holds, then the functions y of Σ depend on $3p - 3 - (3p - \omega) = \omega - 3$ moduli. Otherwise the number of moduli is less than $\omega - 3$.

The treatment of the above outlined questions involves the study of certain continuous systems of curves possessing singularities other than ordinary double points, and may form the subject of another paper. Here, instead, we wish to make some critical remarks concerning a statement in the paper by Severi quoted above, and we also wish to call attention to a

interesting paper,* in which he applies Severi's results to the problem of the moduli of the polygonal curves of genus p.

10. In section 5 of his paper Severi states the following theorem: *On a curve of genus p ($p > 1$) of general moduli the two-fold of any linear series, ∞^1 at least, of order ν (ν arbitrary) is a non-special series.* In stating this theorem Severi remarks that it is an immediate consequence of his general theorem on the index of speciality of the two-fold of a series $g_\nu{}^1$ on a curve of genus p (see this paper, sec. 3). However, it seems to us that, rigorously speaking, this general theorem allows us only to affirm that the two-fold of a *generic* linear series on a curve of genus p of general moduli is non-special. In fact, if the curves of the system H, considered by Severi, are of general moduli, then it is perfectly correct that the two-fold of the $g_\nu{}^1$ on a *generic* curve C of H is a non-special series. But for particular curves of H this series may well become special, and it is not evident at all that the curves of H, for which this happens, are necessarily of *special moduli.* Nothing prevents us from thinking that these particular curves of the system H correspond to *particular* $g_\nu{}^1$'s on curves of genus p of general moduli. The theorem that the two-fold of any series, at least ∞^1, on a curve of genus p of general moduli is non-special, if true (and most probably it is true), must yet be proved.

11. In his quoted paper B. Serge proves that the ν-gonal curves of genus p ($\nu < p/2 + 1$), i. e. the curves of genus p possessing a $g_\nu{}^1$, depend on $\omega - 3 = 2\nu + 2p - 5$ moduli. In view of Severi's theorem, it all amounts to proving the theorem that the index of speciality i of the two-fold of a generic $g_\nu{}^1$ on a general ν-gonal curve of genus p is

$$i = p - 2\nu + 2.$$

Segre's proof of this theorem is somewhat lengthy. We give here a simple proof of this theorem, using the system H considered by Severi. The curves C of H are of order $n + \nu$, of genus p, and possess one fixed ν-fold point P, another fixed n-fold point Q, and $d = (\nu - 1)(n - 1) - p$ variable ordinary double points.

The adjoint curves of order $n + \nu - 3$ of C degenerate into the fixed line PQ and into curves C_1 of order $n + \nu - 4$ with a $(\nu - 2)$-fold point at P and a $(n - 2)$-fold point at Q and passing simply through the double points of C. Let t_1 and t_2 be two generic lines through Q. *We show that there*

* B. Segre, "Sui moduli delle curve poligonali, e sopra un complemento al teorema di esistenza di Riemann," *Mathematische Annalen*, Vol. 100 (1928).

exist exactly $\infty^{p-2\nu+1}$ *curves* C_1 *passing through the* 2ν *intersections (outside of* Q*) of the lines* t_1, t_2 *with* C, *and that hence the index of speciality* i *of the two-fold of the* g_ν^1, *cut out by the lines of the pencil* Q *on the generic curve* C *of* H, *is* $p - 2\nu + 2$.

Let us denote by K the system of the curves C_1, and let δ be the dimension of K. The curves C_1 evidently degenerate into the two lines t_1 and t_2 and into residual curves C_2 of order $n + \nu - 6$, which must possess a $(\nu - 2)$-fold point at P, an $(n - 4)$-fold point at Q and must pass simply through the d double points of C. The system H contains degenerate curves made up of ν arbitrary lines $a_1, a_2, \cdots, a_\nu$, passing through P, and n arbitrary lines $b_1, b_2, \cdots, b_n$ passing through Q. Let D_0 be one of these degenerate curves. As the curve C approaches D_0, the d double points of C tend to d *assigned* double points of D_0 (while the remaining $n + \nu + p - 1$ double points of D_0 have to be considered as *virtually non-existent*), and the curves C_2 will tend to curves Γ_1 of order $n + \nu - 6$, behaving in the same way at P and Q as the curves C_2, and passing simply through the d assigned double points of D_0. The number of arbitrary parameters, on which the limit curves Γ_1 depend, evidently cannot be less than the number δ of arbitrary parameters, on which the curves C_2 depend. Hence if we call Σ_1 the system of the curves Γ_1, and if we denote by δ_1 the dimension of Σ_1, we have

$$\delta_1 \geqq \delta. \tag{16}$$

We proceed to evaluate δ_1. We first agree to choose the $n + \nu + p - 1$ virtually non-existent double points of D_0 as follows: Let us consider the n points in which the line a_1 is met by the lines $b_1, b_2, \cdots, b_n$, and the $3(\nu - 1)$ points in which the lines b_1, b_2, b_3 are met by the lines $a_2, a_3, \cdots, a_\nu$. We have thus considered in all $n + 3(\nu - 1)$ double points of D_0. The remaining

$$n\nu - n - 3(\nu - 1) = (n - 3)(\nu - 1)$$

double points of D_0 are the double points of the curve $D_0' = a_2 + a_3 + \cdots + a_\nu + b_4 + b_5 + \cdots + b_n$. Since by hypothesis $\nu < p/2 + 1$, we have

$$n + \nu + p - 1 > n + 3(\nu - 1),$$

and hence we may assume the $n + 3(\nu - 1)$ double points of D_0 considered above to be virtually non-existent, and choose the remaining

$$(n + \nu + p - 1) - (n + 3\nu - 3) = p - 2\nu + 2 > 0$$

[illegible]
[illegible] $n + \nu + p - 1$ virtually non-existent double points

of D_0, *the d assigned double points of D_0 are double points of D_0'.* For the dimension δ_1 of the system Σ_1 we evidently have

$$\delta_1 \geqq \binom{n+\nu-4}{2} - \binom{\nu-1}{2} - \binom{n-3}{2} - d - 1 = p - 2\nu + 1,$$

where the sign $=$ holds, if the d assigned double points of D_0 impose *d independent conditions* on the curves Γ of order $n+\nu-6$ having a $(\nu-2)$-fold point at P and an $(n-4)$-fold point at Q. To prove that these conditions are indeed independent, we have only to observe that the $(\nu-1)(n-3)$ double points of D_0' impose as many independent conditions on the curves Γ, or, what is the same thing, that the dimension of the system of those curves Γ, which pass through all the double points of D_0', is

$$\binom{n+\nu-4}{2} - \binom{\nu-1}{2} - \binom{n-3}{2} - (\nu-1)(n-3) - 1 = -1,$$

i. e. that such curves Γ do not exist. In fact, a curve Γ, which passes through all the double points of D_0' meets each of the lines $a_2, a_3, \cdots, a_\nu, b_4, \cdots, b_n$ in $(\nu-2)+(n-3)=(\nu-1)+(n-4)=n+\nu-5$ points, and hence necessarily contains each of these lines. But this is impossible, since the order $n+\nu-6$ of Γ is less than the order of the curve $a_2+a_3\cdots+a_\nu+b_4+\cdots+b_n$.

Hence, the $(\nu-1)(n-3)$ double points of D_0' impose independent conditions on the curves Γ, and since the d assigned double points of D_0 are among the double points of D_0', it follows that they also impose independent conditions on the curves Γ. Hence $\delta_1 = p - 2\nu + 1$.

From (16) we deduce that

$$\delta \leqq p - 2\nu + 1,$$

and, since obviously δ cannot be less than $p-2\nu+1$, it follows that

$$\delta = p - 2\nu + 1, \qquad \text{q. e. d.}$$

12. Having thus established the theorem, that the index of speciality of the two-fold of a series g_ν^1 $(\nu < p/2+1)$ on a curve of genus p is *in general* $i = p - 2\nu + 2$, one may ask on how many moduli do the particular ν-gonal curves, for which the above index of speciality is greater than $p-2\nu+2$ depend. In his quoted paper (p. 545) Segre states that the ν-gonal curves for which

$$i = p - 2\nu + 2 + \sigma, \qquad (\sigma \geqq 1)$$

depend on $2p+2\nu-5-\sigma$ moduli. We cannot agree with this statement since we have found that, already in the case $\sigma = 1$, the number of moduli

is less than that given by Segre. It may seem that the above statement is an immediate consequence of Severi's theorem (number of moduli $= 3p - 3 - i$). But it must be emphasized that this theorem says only that, *if i is the index of speciality of the two-fold of a g_ν^1 on the general ν-gonal curve of genus p, then the ν-gonal curves depend on $3p - 3 - i$* moduli. Thus, having found that in general $i = p - 2\nu + 2$, we have deduced that the ν-gonal curves of genus p depend on $2\nu + 2p - 5$ moduli. But Severi's theorem does not say that the *particular* ν-gonal curves, for which the above index of speciality has a given value $i > 2\nu - 2p + 2$, depend on $3p - 3 - i$ moduli.* To give an illustration, let us evaluate the number of the moduli in the case $\sigma = 1$.

LEMMA. *If $(\nu - 1)^2 \geqq p$, there exists an irreducible continuous system $\infty^{4\nu+p-1}$ of plane irreducible curves C of order 2ν and genus p, which possess two fixed infinitely near ν-fold points O, O_1 and $(\nu - 1)^2 - p$ variable ordinary double points.*

We consider the plane curves D of order 2ν possessing a ν-fold point at a fixed point O, and we suppose that O is the origin of ν linear branches having a common fixed tangent t. Under this hypothesis the curves D possess a ν-fold point O_1, infinitely near to O in the direction of the fixed tangent t. The differential condition that a curve of order 2ν should possess a ν-fold point at a fixed point O and a ν-fold point at a point O_1 infinitely near to O on a fixed line through O is expressed by $\nu(\nu + 1)$ linear relations among the coefficients of the equation of the curve, and it is easily seen that these relations are independent.† Hence the curves D form a linear system Σ of dimension

$$\nu(2\nu + 3) - \nu(\nu + 1) = \nu(\nu + 2).$$

The curves of Σ are irreducible, since Σ contains curves D which degenerate into ν arbitrary conics passing through O and touching the line t at O, which shows that the curves of Σ cannot be composed of the curves of a pencil.

Let D_0 be one of the above reducible curves of D, which degenerates into ν conics $c_1, c_2, \cdots, c_\nu$. *We say that the $\nu(\nu - 1)$ double points of D_0 (distinct from O) impose as many independent conditions on the curves D.* In fact let ρ be the dimension of the system Σ' of the curves D passing through the double points of D_0. Then

* Here the question arises whether the family of the said particular ν-gonal curves is irreducible. See Remark 1, p. 171.

† See, F. Enriques and O. Chisini, "Teoria geometrica delle equazioni e delle

$$\rho \geqq \nu(\nu+2) - \nu(\nu-1) = 3\nu, \tag{17}$$

where the sign $=$ holds, if the above conditions are independent.

The curves of Σ' meet the conic c_1 in

$$2\nu + 2(\nu - 1) = 4\nu - 2$$

fixed points, and hence the dimension of the system Σ'' of those curves Σ', which degenerate into the conic c_1 and into a residual curve of order $2(\nu-1)$, is $\geqq \rho - 3$. The curves of Σ'' have a $(\nu-1)$-fold point at O and a $(\nu-1)$-fold point at O_1, and pass through the double points of the reducible curve $c_2 + c_3 + \cdots c_\nu$. Accordingly they meet the conic c_2 in

$$2(\nu-1) + 2(\nu-2) = 4(\nu-1) - 2$$

fixed points, and consequently the dimension of the system Σ''' of those curves of Σ'', which degenerate into the conic c_2 and into a curve of order $2(\nu-2)$, is $\geqq \rho - 6$. This procedure leads after ν steps to the conclusion that the curves of Σ' which degenerate into the conics $c_1, c_2, \cdots, c_\nu$ form a system of dimension $\geqq \rho - 3\nu$. But this last system is of dimension 0, since it contains only one curve, viz., the curve $c_1, + c_2 + \cdots + c_\nu$. Hence,

$$\rho \leqq 3\nu. \tag{18}$$

From (17) and (18) we deduce that

$$\rho = 3\nu.$$

Hence the $\nu(\nu-1)$ double points of the degenerate curve D_0 impose $\nu(\nu-1)$ linearly independent conditions on the curves of the system Σ.

We now consider *as virtually non-existent* $\nu + p - 1$ of the $\nu(\nu-1)$ double points of D_0, so that D_0 becomes a connected curve with two infinitely near assigned ν-fold point at O, O_1 and other $\nu(\nu-1) - (\nu + p - 1) = (\nu-1)^2 - p = d$ assigned double points. The virtual genus of D_0 will be then

$$\binom{2\nu-1}{2} - 2\binom{\nu}{2} - d = p.$$

In order that D_0 should become a connected curve, we may consider, for instance, as virtually non-existent one of the two intersections (distinct from O) of c_1 with each of the conics $c_2, c_3, \cdots, c_\nu$ and other p points chosen arbitrarily from the remaining $(\nu-1)^2$ double points of D_0.* In the linear

* For the concepts which we use here, and also for the method of "analytical regions" ("falde analitiche," "analytische Mantel") see Severi, quoted treatise, *Anhang F*. See also quoted treatise by *Enriques-Chisini*, Vol. 3, Chap. 3, § 33.

space $S_{\nu(\nu+2)}$ of dimension $\nu(\nu+2)$, the points of which represent the curves of the system Σ, the curves of Σ possessing one double point in the neighborhood of one of the assigned double points of D_0 are represented by a linear analytical region Φ_i $(i = 1, 2, \cdots d)$ having its origin at D_0. Since the assigned d double points of D_0 impose linearly independent conditions on the curves of Σ, the d tangent spaces of the regions Φ_i at D_0 are linearly independent. We deduce that the Φ_i have in common an analytical region Φ of dimension

$$\nu(\nu+2) - d = 4\nu + p - 1.$$

The curves of Φ possess $d = (\nu - 1)^2 - p$ double points in the respective neighborhoods of the assigned double points of D_0, *and hence D_0 belongs to a complete continuous system H of curves C of order 2ν, possessing the two infinitely near ν-fold points O, O_1 and* $(\nu - 1)^2 - p$ ordinary double points.

That the generic curve C of H possesses only ordinary double points (distinct from O and O_1), follows from the fact that this holds for the particular curve D_0 of H. Moreover the generic curve C of H does not have more than $(\nu - 1)^2 - p$ double points, because, if it had $d + 1$ double points, then, as C approaches D_0, the double points of C would tend to $d + 1$ double points of D_0. Since these $d + 1$ double points of D_0 would also impose independent conditions on the curves of Σ, it would follow that the dimension of H is $4\nu + p - 2$ and not $4\nu + p - 1$ as above. Hence, the generic curve of H has exactly $(\nu - 1)^2 - p$ double points (distinct from O and O_1) and therefore is of genus p.

It remains to prove that the curves of H *are irreducible.* This follows from the fact that on the degenerate curve D_0 any two points can be joined by a path which does not meet any of the $(\nu - 1)^2 - p$ assigned double points, and that therefore also on the generic curve C of H it is possible to connect any two points by a continuous path without meeting any of its double points. The Lemma is thus proved.

On a generic curve C of the system H the lines of the pencil O cut out a linear involution g_ν^1, and hence C is a ν-gonal curve of genus p. The ∞^3 conics C passing through O and touching the fixed tangent t at O (and, in particular, the ∞^2 lines of the plane) cut out on C sets of 2ν points belonging to the two-fold of the series g_ν^1. We prove that, *if $\nu \leq p/2 + 1$, the series $g_{2\nu}^2$ cut out by the above conics is complete, or that the index of speciality i of the two-fold of the g_ν^1 is $p - 2\nu + 3$.*

To find i, we have to find the number of linearly independent adjoint [illegible]

[illegible] the curve C. The adjoint [illegible]

of order $2\nu-3$ of C possess two infinitely near $(\nu-1)$-fold points at O, O_1, and hence degenerate into the fixed line t and into residual curves of order $2\nu-4$ possessing two infinitely near $(\nu-2)$-fold points at O, O_1, and passing simply through the remaining $(\nu-1)^2-p$ double points of C. The adjoint curves of order $2\nu-3$ of C, which pass through the above 2ν intersections of the conic c with the curve C, degenerate into the fixed line t, into the conic c and into residual curves C_1 of order $2\nu-6$ possessing two infinitely near $(\nu-3)$-fold points at O, O_1 and passing simply through the remaining $(\nu-1)^2-p$ double points of C. Since the dimension of the system of the curves C_1 is, by hypothesis, $i-1$, we have

$$(19) \qquad i \geqq \binom{2\nu-4}{2} - 2\binom{\nu-2}{2} - (\nu-1)^2 + p = p - 2\nu + 3.$$

Now, let the curve C tend to the degenerate curve $D_0 = c_1 + c_2 + \cdots + c_\nu$. The curves C_1 will become in the limit, curves Γ of order $2\nu-6$, behaving in the same way as the curves C_1 at the points O, O_1 and passing simply through the $(\nu-1)^2-p$ assigned double points of D_0. Hence, denoting by r the dimension of the complete linear system defined by the base points of the curves Γ, we will have

$$(20) \qquad r \geqq i - 1.$$

In order to evaluate r, we first agree to choose the $(\nu-1)^2-p$ assigned double points of D_0 as follows: We consider the following $3(\nu-1)$ double points of D_0: (a) The $2(\nu-1)$ points at which the conic c_1 is met by the conics $c_2, c_3, \cdots, c_\nu$; (b) the two intersections of the conics c_2, c_3; (c) one of the two points at which c_2 is met by each conic $c_4, \cdots, c_\nu$. If the above $3(\nu-1)$ double points of D_0 are considered as virtually non-existent, D_0 becomes a connected curve. Since, by hypothesis, $\nu \leqq p/2 + 1$, and hence $\nu + p - 1 \geqq 3(\nu-1)$, we may assume the considered $3(\nu-1)$ double points of D_0 to be virtually non-existent, and choose arbitrarily the remaining

$$\nu + p - 1 - 3(\nu-1) = p - 2\nu + 2 \geqq 0$$

virtually non-existent double points of D_0 from the remaining

$$\nu(\nu-1) - 3(\nu-1) = (\nu-1)(\nu-3)$$

double points of D_0. Let us call A this last set of $(\nu-1)(\nu-3)$ double points of D_0.

The $(\nu-1)(\nu-3)$ *points* A *impose as many independent conditions on the curves* E *of order* $2\nu-6$ *possessing two* $(\nu-3)$-fold points at O, O_1.

In fact, let r' be the dimension of the system K of those curves E which pass through the points A. Then

$$r' \geqq \binom{2\nu-4}{2} - 2\binom{\nu-2}{2} - (\nu-1)(\nu-3) - 1 = 0,$$

where $r' = 0$, if and only if the points A impose $(\nu-1)(\nu-3)$ independent conditions on the curves E. Now let us observe that each of the conics $c_4, c_5, \cdots, c_\nu$ contains $2\nu - 5$ points A and that consequently each curve E of K must contain each of the above conics, since each curve E of K has already, outside of the points A, $2\nu - 6$ coincident intersections with E at O. It follows that the system K contains only one curve, viz. the degenerate curve $c_4 + c_5 + \cdots + c_\nu$, and hence $r' = 0$.

Since the $(\nu-1)^2 - p$ assigned double points of D_0 are all chosen from the points A, they also impose independent conditions on the curves E. It follows that

$$r = \binom{2\nu-4}{2} - 2\binom{\nu-2}{2} - (\nu-1)^2 + p - 1 = p - 2\nu + 2,$$

and hence, from (20),

(21) $$i \leqq p - 2\nu + 3.$$

Comparing (21) with (19) we deduce that

$$i = p - 2\nu + 3, \qquad \text{q. e. d.}$$

Let C_0 be a generic curve of H, and let us consider the $\sigma = 2\nu + 2p - 2$ tangents λ to C_0 through O. We consider the system of curves C of H, which touch the tangents λ, and we denote by H_0 that irreducible part of this system which contains C_0. It is easily seen that the sets of characteristic series of H in C_0 are sets of the two-fold of the series g_ν^1. On the other hand, by a theorem, due to Severi, on continuous systems of curves constrained to touch a fixed set of lines of a pencil, which we have used at an earlier stage of this paper (see sec. 8), and which applies without any modification to the system H_0, we deduce that the characteristic series of H_0 on C_0 is *the complete series* $g_{2\nu}^8$.

It follows that the σ tangents λ depend only on $4\nu + p - 1 - 4 = 4\nu + p - 5$ arbitrary parameters, and that hence the curves C of H depend *at most* on $4\nu + p - 8$ moduli. It will be noted that if $\nu < p/2 + 1$, the derived upper limit $4\nu + p - 8$ of moduli is certainly less than the number $2\nu + 2p - 6$ evaluated according to Segre's formula. *We have thus estab-*

genus p, for which the two-fold of the series g_ν^1 is a special series of index of speciality $p - 2\nu + 3$, and which depend at most on $4\nu + p - 8$ moduli. We observe that, in order to prove that the curves of H depend exactly on $4\nu + p - 8$ moduli, it is necessary to prove that the generic curve of H does not possess infinite linear involutions g_ν^1.

Remark 1. If a ν-gonal curve of genus p contains a g_ν^1 such that the two-fold of the g_ν^1 is a $g_{2\nu}^3$ (at least), the curve can be transformed into a plane curve of order 2ν possessing two fixed infinitely near ν-fold points and other singularities equivalent to $(\nu - 1)^2 - p$ double points, *provided the $g_{2\nu}^3$ is not composed of an involution.* In order to prove that the ν-gonal curves for which the above $g_{2\nu}^3$ is not composed of an involution form an irreducible system, it would be necessary to prove that the plane curves of order 2ν and of genus $p \leqq (\nu - 1)^2$, possessing two fixed infinitely near ν-fold points, form a unique irreducible system, which then coincides with the system H considered above.

Under the hypothesis that the $g_{2\nu}^3$ is composed of an involution, it can be easily verified that the corresponding ν-gonal curves may form a complete irreducible system of dimension $> 4\nu + p - 8$ and hence distinct from H.

Remark 2. We suggest here a formula for the number of moduli of the ν-gonal curves of genus p possessing a g_ν^1, such that the two-fold series of the g_ν^1 is of index of speciality $i = p - 2\nu + 2 + \sigma$ (always under the hypothesis that the corresponding complete series $g_{2\nu}^{\sigma+2}$ is not composed of an involution). This formula is as follows:

$$\begin{aligned}\text{number of moduli} &= 2\nu + 2p - 5 - \sigma(p - 2\nu + 2 + \sigma)\\ &= 2\nu + 2p - 5 - \sigma i.\end{aligned}$$

The above formula gives, for $\sigma = 1$, the derived number $4\nu + p - 8$ above, and for any $\sigma > 0$ is suggested by the following hyperspacial considerations, relative to the canonical curve of genus p in a space S_{p-1}. Assuming that the canonical curve C is a ν-gonal curve, each of the ∞^1 sets Γ_ν of the g_ν^1 belongs to a space $S_{\nu-2}$. The index of speciality of the two-fold of the series g_ν^1 being $p - 2\nu + 2 + \sigma$, any two of the spaces $S_{\nu-2}$ meet in a space $S_{\sigma-1}$. We know that if C is a generic ν-gonal curve, then $\sigma = 0$, i. e. the spaces $S_{\nu-2}$ do not meet one another. It is also easily seen that the number of conditions, which express the fact that two spaces $S_{\nu-2}$, given in a S_{p-1}, meet in a $S_{\sigma-1}$, is $\sigma(p - 2\nu + 2 + \sigma) = \sigma i$. The above formula is derived under the hypothesis that to the above number of incidence conditions corresponds an equal number of independent conditions on the moduli of the curve C.

THE JOHNS HOPKINS UNIVERSITY.

On Self-Adjoint Ordinary Differential Equations of the Fourth Order.*

BY WILLIAM M. WHYBURN †

The work of Sturm, Bôcher, and others on real second order differential equations forms an important chapter in mathematical analysis. Birkhoff ‡ obtained many important results for third order differential systems and introduced a new and valuable method of proving comparison theorems for such systems. C. N. Reynolds § used Birkhoff's methods to study equations of the n^{th} order. Hille ‖ studied the second order equation in the complex domain and discovered many important facts concerning the distribution of the zeros of the solution in the complex plane. Davidoglou ¶ gave a detailed treatment of a special fourth order self-adjoint equation that arose in connection with a study of the lateral vibrations of a rod. If certain factors, such as the rotational inertias, are taken into account in the problem of the vibrating rod, the equation that must be studied is the general self-adjoint fourth order equation rather than the special equation treated by Davidoglou.** H. T. Davis †† has recently considered this equation.

It is the intention of the present paper to study the general self-adjoint equation of the fourth order

$$\frac{d^2}{dx^2}\left[K(x)\frac{d^2u}{dx^2}\right]+\frac{d}{dx}\left[L(x)\frac{du}{dx}\right]+G(x)u=0 \tag{1}$$

and the second order equation

$$\frac{d^2u}{dx^2}+[p_1(x)+i\,p_2(x)]\frac{du}{dx}+[q_1(x)+i\,q_2(x)]u=0, \tag{2}$$

where K, L, G, p_1, p_2, q_1, q_2 are real functions of the real variable x and i is

* Presented to the American Mathematical Society, February 25, 1928.

† National Research Fellow in Mathematics.

‡ *Annals of Mathematics*, Ser. 2, Vol. [illegible] (1910-11), pp. 103-127.

§ *Transactions of American Mathematical Society*, Vol. 22 (1921), pp. 220-229.

‖ *Transactions of American Mathematical Society*, Vol. 23 (1922), pp. 350-385.

¶ *Annales de l'Ecole Normale Superieure* (3), Vol. 22, pp. 539 ff.

** Cf. Lord Rayleigh: *Theory of Sound*, London, 1894, Vol. I, Chapter VIII.

†† [illegible]

the imaginary symbol. This study is accomplished with the aid of a pair of equations of the second order.* The paper treats system (3) for the cases $qr < 0$ on the given interval and $qr > 0$ on that interval. Extensive use is made of the associated second order equation (6) and in many cases the properties of solutions of system (3) are derived from, or related to, similar properties of solutions of (6). Use is made of a combination $V(x)$ of y and z and by means of the properties of this function and transformations involving solutions of equation (6) separation theorems are established for the zeros of $y(x)$ and $z(x)$ in the case $qr < 0$. In the presence of additional restrictions on p, q, and r, it is shown that for every solution of (3) $y(x)$ and $z(x)$ oscillate infinitely often on the real axis and the integral curve in the yz-plane is shown to be essentially *spiriliform.* Related and equivalent integral equations are used in Part III to develop comparison theorems on "regular" intervals for system (3) in the case $qr > 0$. The paper leaves the cases that arise when qr changes sign on the interval wholly untouched. The case $qr < 0$ is emphasized because it is the one which arises in a simultaneous treatment of equations (1) and (2). This point is brought out more clearly near the end of Part I.

I. TRANSFORMATION OF THE EQUATIONS.

We follow the well established custom of stating that $f(x)$ is of class $C^{(j)}$ on X: $a \leqq x \leqq b$ provided $f(x)$ together with its first j derivatives is continuous on X.

THEOREM I. *If, in equation* (1), *K, L'', G are continuous and $K \neq 0$ on X, equation* (1) *can be written in the form*

$$(3) \qquad \begin{aligned} &(a)\ \ y'' + p(x)y = q(x)z \\ &(b)\ \ z'' + p(x)z = r(x)y, \end{aligned}$$

where p, q, and r are continuous on X and furthermore, $q(x) \neq 0$ and p/q is of class C'' on X. Conversely, if p, q, and r are continuous on x and $q \neq 0$, p/q is of class C'' on X, then system (3) *is equivalent to equation* (1) *where $K \neq 0$ and K, L'', and G are continuous on X.*

Proof. Let $q = 1/K$, $p = L/2K$, $r = L''/2 + L^2/4K - G$. Solving these equations for K, L, and G *yields*

$$K = 1/q,\ L = 2p/q,\ G = (p/q)'' + p^2/q - r$$

* The author is indebted to Prof. G. D. Birkhoff for many valuable suggestions that have furthered the writing of this paper.

and it becomes evident that p, q, and r have the desired properties of continuity, etc.

Substitution into equation (1) yields

$$[u''/q]'' + 2[pu'/q]' + u(p/q)'' + up^2/q - ru = 0$$
$$[u''/q + pu/q]'' + p[u''/q + pu/q] = ru.$$

Let $u = y$ and $z = u''/q + pu/q$. Simplification after this substitution yields system (3).

To prove the converse, we divide equation (3) (a) by q and compute z'' from the resulting equation. Upon substituting this value into equation (3) (b) and simplifying, one obtains equation (1).

In equation (2) let p_1 and p_2 be of class C' on X while q_1 and q_2 are continuous on that interval. Let $u = W \exp\left[-\tfrac{1}{2}\int_a^x (p_1 + ip_2)\,dt\right]$, $Q = p + iq$, where $p = q_1 - p_1^2/4 + p_2^2/4 - p_1'/2$ and $q = q_2 - p_1p_2/2 - p_2'/2$. Equation (2) becomes

$$W'' + Q(x)W = 0, \tag{4}$$

where Q is continuous on X.

The ordinary existence theorem* insures the existence of a unique solution † of (4) that is of class C'' on X and such that if $x = c$ is any point of X and $\alpha, \beta, \gamma, \delta$, are arbitrarily assigned real constants, this solution takes on the initial values $W(c) = \alpha + \beta i$, $W'(c) = \gamma + \delta i$. A similar theorem applied to the system

$$y'' + p(x) = q(x)z, \qquad z'' + p(x)z = -q(x)y \tag{5}$$

shows the existence of a unique pair of functions of class C'' on X that satisfies (5) on X and takes on the initial values $y(c) = \alpha$, $y'(c) = \gamma$, $z(c) = \beta$, $z'(c) = \delta$. Let $W = y + iz$. Direct substitution shows that W is a solution of (4) that assumes the given initial values. It follows from the uniqueness of such solutions that (4) with its initial conditions is equivalent to (5) with the corresponding initial conditions. It also follows that any solution of (4) that is of class C'' on X can be written in the form $y + iz$, where y and z are real functions of the real variable x.

Thus we are led to a consideration of the pair of real equations (3). Since the transformation of Theorem 1 carries the non-singular equation (1)

---

* Cf. Bôcher, *Leçons sur les Méthodes de Sturm*. Paris (1917), page 1.

† [illegible]

into a system (3) where $q(x)$ does not vanish on X while equation (2) yields a system (3) with $r(x) = -q(x)$, we are led to a study of (3) where $q(x)r(x) < 0$ in order to obtain results that carry over to both equation (1) and equation (2).

Let x be regarded as a parameter, then the pair of functions $y(x)$, $z(x)$ given by system (3) describes a curve in the yz-plane as x ranges over the set of values X. This curve is called the *integral* curve of (3) in the yz-plane and its shape is evidently dependent upon the initial values as well as the coefficients of the system. We propose to study this curve. We use the notation (R, θ) for the polar co-ordinates of a point on this curve and hence $R(x)^2 = y(x)^2 + z(x)^2$, $\theta(x) = \tan^{-1}(z/y)$.

The following dynamical interpretation may be given for system (3) following a similar one given by Hille[*] for system (5). If x is regarded as the time, then the integral curve of (3) in the yz-plane is the path of motion of a particle that is acted upon by a force of magnitude $|pR|$ along the radius vector, a force of magnitude $|qR|$ perpendicular to the radius vector, and a force of magnitude $|(r - q)y|$ parallel to the z-axis. Other distributions of the forces may be used. In the light of this interpretation, the hypotheses and conclusions of later theorems in the paper have a definite physical significance.

II. SYSTEM (3) WITH $q(x)r(x) < 0$.

Throughout this part we suppose that p, q, and r are continuous on the interval under consideration and in addition that $qr < 0$ on this interval.[†] Let

$$V(x) = V[y(x), z(x)] = y'(x)z(x) - z'(x)y(x).$$

If equation (3) (b) is multiplied by $y(x)$ and this result subtracted from the result of multiplying (3) (a) by $z(x)$, one obtains $V'(x) = q(x)z(x)^2 - r(x)y(x)$.[2] Upon integrating from $x = c$,

$$V(x) = V(c) + \int_c^x [q(t)z(t)^2 - r(t)y(t)^2]\,dt.$$

Since both y and z cannot vanish over an everywhere dense set of points in any subinterval of X without vanishing identically on X, we have that $V(x)$

* *Loc. cit.*, page 359.

† A number of the results of the section are valid if $r(x)$ is allowed to vanish (without changing signs) over a nul set of points.

is an increasing or decreasing function of x on X according as $q(x) > 0$ on X or $q(x) < 0$ on that interval.*

THEOREM II. *The zeros of $y(x)$ and $z(x)$ separate each other on X with the possible exception of a neighborhood of one point at which $V(x) = 0$.*

Proof. Calculate $\theta'(x)$. $\theta'(x) = -V(x)/R(x)^2$, unless $R(x) = 0$. Since $V(x)$ can vanish at most once on X and since V vanishes with R, it follows that with the possible exception of one point, θ' is given by the above formula and hence can vanish or fail to exist for at most one value of x on X. Call this exceptional value c if it exists. On $a \leqq x < c$, $\theta'(x)$ is different from zero and of one sign so that θ changes continuously in one sense and the integral curve must cross the y and z axes alternately. Similarly, on $c < x \leqq b$, θ must change continuously in one sense and the zeros of y and z separate each other on this interval.

The above proof shows that under the hypotheses of this section θ can change its sense of variation at most once and the integral curve of (3) cannot pass through the origin of the yz co-ordinate system more than once.

Since $R^2\theta'$ is the rate at which area is swept out by the radius vector and the integral curve and since this quantity is equal to $-V(x)$, we have that the rate of sweeping out area increases or decreases with increase of x according as $q(x)$ is negative or positive.

THEOREM III. *If the closed interval X is of finite length, neither y nor z can vanish more than a finite number of times on X.*

Proof. Assume that y vanishes infinitely often on X and let $x = c$ be a limit point of its zeros on X. It follows from Theorem II that $x = c$ is also a limit point of the zeros of z. An application of Rolle's theorem shows that $x = c$ must be a limit point of the zeros of both $y'(x)$ and $z'(x)$. Since $y(x)$ and $z(x)$ are of class C'' on X, we have $y(c) = z(c) = y'(c) = z'(c) = 0$. It follows from the fundamental existence theorem that $y(x) \equiv z(x) \equiv 0$ on X. This, however, contradicts the fact that we are working with non-identically vanishing solutions of (3).

1°. *Case where the solutions of* (6) *oscillate on X.* Consider the second order equation

$$W'' + p(x)W = 0, \tag{6}$$

where $p(x)$ is the function that appears in system (3). It is well known †

[illegible]

[illegible]

that the zeros of any two linearly independent solutions of (6) separate each other on X.

THEOREM IV. *If there exists a solution of* (6) *that vanishes twice on* X, *then the product* $y(x)z(x)$ *must vanish at least once on* X *for every solution* $y(x)$, $z(x)$ *of system* (3).

Proof. Suppose that y, z is a solution of (3) such that yz does not vanish on X. Let $u(x)$ and $v(x)$ be the unique functions determined by the second order systems

$$(7) \quad [y^2u']' + qyz\,u = 0, \qquad u(a) = W(a)/y(a), \qquad u'(a) = [W/y]'_{x=a},$$
$$(8) \quad [z^2v']' + ryz\,v = 0, \qquad v(a) = W(a)/z(a), \qquad v'(a) = [W/z]'_{x=a},$$

where $W(x)$ is the solution of (6) that vanishes twice on X. The functions $u(x)y(x)$ and $v(x)z(x)$ are each solutions of (6) since

$$[uy]'' = [y^2u']'/y + uy'' = [-qz + y'']u = -pyu,$$

and similarly for vz. Furthermore, $u(a)y(a) = W(a)$, $[uy]'_{x=a} = W'(a)$, $v(a)z(a) = W(a)$, $[vz]'_{x=a} = W'(a)$. It follows from the uniqueness of $W(x)$ that

$$(9) \qquad W(x) \equiv u(x)y(x) \equiv v(x)z(x).$$

By virtue of the hypotheses on q and r we have that either $qyz < 0$ or $ryz < 0$ on X. It follows from well-known theorems* for second order equations of type (7) and (8) that either $u(x)$ or $v(x)$ cannot vanish twice on X according as $qyz < 0$ or $ryz < 0$ on that interval. Since neither y nor z vanishes on X and W vanishes twice, it follows from (9) that both u and v must vanish twice on X. This contradiction yields our theorem.

COROLLARY I. *If* (6) *has a solution that vanishes four times on* X, *then* y *and* z *each have at least one zero on* X *and the product* yz *vanishes at least three times on that interval.*†

2°. *The case* $p(x) \equiv 0$. Let $p(x)$ be identically zero and consider the resulting pair of equations. In the dynamical interpretation of Part I this means that the force along the radius vector is absent.

---

* Cf. Bôcher, *loc. cit.*, page 51.

† Theorem II shows that the number of zeros of y and z on X can differ by at most two. The corollary follows from this fact and the observation that yz must vanish at least three times.

THEOREM V. *Let the coefficients of system* (3) *be continuous on the interval* X: $a \leqq x < \infty$ *and let* $p(x) \equiv 0$ *on this interval, then every solution of* (3) *such that* $V(a)q(a) \geqq 0$ *will oscillate infinitely often on* X *provided the Riemann integrals* $\int_a^b tq(t)\,dt$, $\int_a^b tr(t)\,dt$ *become infinite with* b. *That is to say, the integral curve in the yz-plane will circulate the origin infinitely often as* x *ranges over* X.

Proof. From the separation of the zeros of y and z, it follows that if one of these vanishes only a finite number of times the other will also. Assume that there exists a solution $y(x)$, $z(x)$ of (3) such that $V(a)q(a) \geqq 0$ and such that y and z vanish only a finite number of times on X. Let $x = d$ be chosen on X so that $d > a$ and neither y nor z vanishes on D: $d \leqq x < \infty$. For the sake of definiteness, we suppose $q > 0$ on X since the other case can be obtained from this by merely interchanging the roles of y and z in the arguments. Since $V(x)$ is an increasing function on X, it follows that $V(x) > 0$ on D and hence either: (A) $y'(d)z(d) > 0$, or (B) $y(d)z'(d) < 0$.

Consider (A). From the differential equations,

$$y'(x) = y'(d) + \int_d^x q(t)z(t)\,dt,$$

$$z'(x) = z'(d) + \int_d^x r(t)y(t)\,dt.$$

Since $y'(d)z(d) > 0$, it follows from the above formulas that $y'(x)$ is a positive increasing function or a negative decreasing function on D according as $z(x)$ is positive or negative on D. Since the same arguments treat both cases, we suppose that $z(x) > 0$ on D. We may choose a point $x = h$ on D so that $\int_d^h y'(t)\,dt > |\, y(d) \,|$ and we then have for every $x > h$, $y(x) > \int_h^x y'(t)\,dt > y'(h)(x-h)$. Now $z'(x) = z'(h) + \int_h^x r(t)y(t)\,dt > z'(h) + y'(d) \int_h^x r(t)(t-h)\,dt$. The right hand side of this inequality diverges to $-\infty$ as x increases indefinitely and hence $z(x) = z(h) + \int_h^x z'(t)\,dt$ becomes negatively infinite as x becomes infinite on D. But $z(x)$ is continuous and $z(d) > 0$. Hence $z(x)$ must vanish on D thereby contradicting our hypothesis that $z(x) \neq 0$ on D.

The arguments for (B) follow the same lines as for (A). The roles of y and z are interchanged in the proof and a contradiction is reached by

THEOREM VI. *If the hypotheses of Theorem V are met with the exception that the integrals* $\int_a^\infty q(t)\,dt$, $\int_a^\infty r(t)\,dt$ *are required to diverge instead of the two integrals there given, then* EVERY *solution of* (3) *either oscillates infinitely often on* X *or there exists a point* $x = h$ *on* X *such that the quantity* $R^2 = y^2 + z^2$ *approaches zero steadily on* H: $h \leqq x < \infty$.

Proof. If for any x, $x = c$, on X, we have $V(c)q(c) \geqq 0$, then Theorem V applies to yield the present theorem. It remains to treat the case $V(x)q(x) < 0$ for every x on X and neither y nor z vanishes infinitely often on X. Let d be chosen so that $d > a$ and neither y nor z vanishes on D: $d \leqq x < \infty$. It follows from the differential equations that y'' and z'' cannot vanish on this interval and hence by Rolle's theorem that y' and z' can each vanish at most once on D. Let h be chosen so that none of the quantities y, z, y', z' vanish on H: $h \leqq x < \infty$. Since $V(x)q(x)$ is negative on H, it follows from the hypotheses on the integrals of q and r over X and the formula $V(x) = V(h) + \int_h^x (qz^2 - ry^2)\,dt$ that neither y^2 nor z^2 can be bounded from zero on H. Hence $yy' < 0$, $zz' < 0$ on H and $(R^2)' = 2(yy' + zz') < 0$. Furthermore, both y^2 and z^2 approach zero and hence R^2 approaches zero steadily as x becomes infinite on H.

3°. *The case* $p(x) < 0$, $p'(x) \leqq 0$ *on* X. In this section we suppose that $p'(x)$ exists on X: $a \leqq x < \infty$ and is continuous along with q and r on this interval. Furthermore, we require that $p(x) < 0$ and $p'(x) \leqq 0$ on X. In the dynamical system of Part I this requirement means that the force along the radius vector is directed away from the origin and that its magnitude can decrease only if R decreases.

THEOREM VII. *If there exists a positive constant* J *such that* $|q|/(-p)^{1/2} \geqq J$, $|r|/(-p)^{1/2} \geqq J$ on X, *then every solution of* (3) *such that* $V(a)q(a) \geqq 0$ *oscillates infinitely often on* X.

For the sake of definiteness, we suppose that $q(x)$ is positive on X since the other case is obtained from this by simply interchanging the roles of y and z in the proofs.

LEMMA I. *If* $w(x)$ *is the solution of* (6) *such that* $w(c) = 1$, $w'(c) = 0$, *where* $x = c$ *is any fixed point of* X, *then* $q(x)w(x)/w'(x) > q(x)/[-p(x)]^{1/2}$, $-r(x)w(x)/w'(x) > -r(x)/[-p(x)]^{1/2}$ on D: $d \leqq x < \infty$, *where* d *is any number that is greater than* c.

Proof of Lemma I. Since $p(x) < 0$ on X, ww' can vanish at most once on this interval. It follows from $w'(x) = -\int_c^{\infty} p(t)w(t)\,dt$ that w and w' are positive on D. Multiplying (6) by $w'(x)/(-p)$ and integrating from $t = c$ to $t = x$, we get $w(x)^2 = 1 + 2\int_c^{\infty} w''(t)w'(t)\,dt/[-p(t)]$. Integrating this by parts

$$w(x)^2 = 1 + w'(x)^2/[-p(x)] + \int_c^{\infty} (1/p)'w'(t)^2 dt.$$

Since $p' \leqq 0$ and $w' > 0$ on D, we have $w(x)^2 > w'(x)^2/[-p(x)]$ or $w/w' > 1/(-p)^{1/2}$ and multiplication of this inequality by q and $-r$, respectively, yields the desired inequalities.

LEMMA II. *If $x = b$ is any point of D and K and H are constants, K being positive, then*

$$\lim_{x \to \infty} \int_b^{\infty} K\Big[\int_b^s f(t)w(t)^2 dt + H\Big] ds/w(s)^2 = +\infty,$$

where $f(t)$ is identical with either $q(t)$ or $-r(t)$ and D, c, w have the same significance as in Lemma I.

Proof of Lemma II. Since $\lim_{x\to\infty} w(x)^2 = +\infty$, let g be so chosen that for every x on $G\colon g \leqq x < \infty$, $w(b)^2/w(x)^2 < \epsilon$, where $0 < \epsilon < 1$ and $g > b$. On $C\colon c \leqq x < \infty$, we have $w'' = -pw \geqq -p(a) > k > 0$, hence $w'(x) \geqq k(x - c)$ and $w(x) \geqq k(x-c)^2/2$. On D, $|H|/w^2 < 4|H|/k^2(x-c)^4$ and hence $\int_b^{\infty}[H/w(s)^2]ds$ converges as x becomes infinite. The proof of Lemma II will be completed if we show that

$$I = \int_b^{\infty} \Big[\int_b^s w(t)^2 f(t)\,dt/w(s)^2\Big] ds$$

diverges as x becomes infinite. Let x be any point of G and apply Cauchey's theorem* to the integrand of I.

$$\int_b^{x} w(t)^2 f(t)\,dt/w(x)^2 = f(s)w(s)[1 - w(b)^2/w(x)^2]/2w'(s),\ b < s < x,$$
$$> (1-\epsilon)q(s)w(s)/2w'(s) > (1-\epsilon)q(s)/2[-p(s)]^{1/2}$$

* Cf. Pierpont, *Theory of Functions of a Real Variable*, Boston (1905), Vol. I, page 330. This theorem states that if $f(t)$ and $g(t)$ are continuous on $b \leqq t \leqq x$ while [illegible] where [illegible]

by Lemma I. Hence $I > \int_g^x (1-\epsilon)J dt/2 = (1-\epsilon)(J/2)(x-g)$ and becomes infinite with x. This fact combined with the above proof that $\int_b^x H ds/w(s)^2$ converges as x becomes infinite on X yields Lemma II.

Proof of Theorem VII. Make the change of variable $y = wY$, $z = wZ$ on C: $c \leqq x < \infty$. Since $w(x)$ does not vanish on this interval and since $\theta = \tan^{-1}(z/y) = \tan^{-1}(Z/Y)$ is an invariant under this transformation, we will have proved our theorem if we show that $Y(x)$ and $Z(x)$ oscillate infinitely often on C. Furthermore,

$$U(x) = Y(x)Z'(x) - Z(x)Y'(x) = (yz' - zy')/w^2 = V(x)/w^2$$

and hence $U(x)$ is of the same sign as $V(x)$. The differential equations (3) become

$$Y'' + 2(w'/w)Y' = qZ, \qquad Z'' + 2(w'/w)Z' = rY, \tag{7}$$

which when solved by the method of variation of parameters yields

$$(8) \quad \begin{aligned} &\text{(a)} \quad Y'(x) = \left[\int_b^x w(t)^2 q(t) Z(t)\,dt + w(b)^2 Y'(b)\right]/w(x)^2 \\ &\text{(b)} \quad Z'(x) = \left[\int_b^x w(t)^2 r(t) Y(t)\,dt + w(b)^2 Z'(b)\right]/w(x)^2, \end{aligned}$$

where $x = b$ is any point of C.

Now assume that Y and Z vanish only a finite number of times on C and let D: $d \leqq x < \infty$ be chosen as a subinterval of C on which neither Y nor Z vanishes. Since $V(d) > 0$, it follows that $U(d)$ is positive. From $U(d) > 0$ follows that either: (A) $Y'(d)Z(d) > 0$, or (B) $Z'(d)Y(d) < 0$.

Consider (A). For the sake of definiteness, we assume $Y'(d) > 0$, $Z(d) > 0$ although a treatment of the case $Y'(d) < 0$, $Z(d) < 0$ merely involves the replacement of Y, Z, Y', and Z' by their negatives in the following proof. It is convenient to treat (A) in three sub-cases: *Case* A_1, $Y(d) > 0$; *Case* A_2, $Y(d) < 0$, $Z'(d) \geqq 0$; *Case* A_3, $Y(d) < 0$, $Z'(d) < 0$.

Case A_1. Since $Z(x)$ and $Y'(d)$ are positive and $Z(x)$ does not vanish on D, we have from (8) (a) that $Y'(x)$ is positive on D. Hence $Y(x)$ is a positive increasing function on D and $Y(x) \geqq Y(d)$ on that interval. From (8) (b), we have

$$\begin{aligned} -Z'(x) &= \left[-\int_d^x r(t)w(t)^2 Y(t)\,dt + H\right]/w(x)^2 \\ &\geqq Y(d)\left[-\int_d^x rw^2 dt + H\right]/w^2, \end{aligned}$$

where $H = -w(d)^2 Z'(d)$. Hence

$$-Z(x) = -Z(d) + \int_d^x [-Z'(t)]\,dt \geqq -Z(d) + \int_d^x [Y(d) \int_d^s -r(t)w(t)^2 dt + H]\,ds/w(s)^2.$$

By Lemma II, the integral on the right hand side of this inequality becomes positively infinite with x and hence $\lim_{x\to\infty} Z = -\infty$. From the continuity of $Z(x)$ and $Z(d) > 0$ follows that $Z(x)$ must vanish at least once on D, thereby contradicting our hypothesis that $Z(x) \neq 0$ on D.

Case A_2. From $Z'(d) \geqq 0$, $Y(d) < 0$, $r(x) < 0$, we get that $Y(x)$ is negative on D and hence from (8) (b) we have that $Z'(x) \geqq 0$ on D. Hence $Z(x)$ is a positive non-decreasing function on D and $Z(x) \geqq Z(d) > K > 0$. From equation (8) (a), we have

$$Y'(x) \geqq [K \int_d^x w(t)^2 q(t)\,dt + H]/w(x)^2, \qquad H = w(d)^2 Y'(d),$$
$$Y(x) \geqq Y(d) + \int_d^x [K \int_d^s w(t)^2 q(t)\,dt + H]\,ds/w(s)^2.$$

Applying Lemma II to the integral on the right hand side, we get $\lim_{x\to\infty} Y(x) = +\infty$ and hence $Y(x)$ must vanish at least once on D, since $Y(d) < 0$. This contradicts our hypothesis that $Y(x)$ does not vanish on D.

Case A_3. An examination of equation (8) (b) shows that if for any value, $x = b$, of x on D, we have $Z'(b) \geqq 0$, then for $x > b$, $Z'(x) > 0$ and the argument of *Case* A_3 applies to give a contradiction. Hence $Z'(x) < 0$ on D.

We prove that neither $Y(x)$ nor $Z(x)$ can vanish on $c \leqq x \leqq d$. Assume that one or the other of these functions vanishes on C and let $x = g$ be the last zero of YZ on C. Since $Z'(x)$ is an increasing function of x on G: $g < x < \infty$, we have $Z'(g) < 0$. If $Y(g) = 0$, then $U(g) = Y'(g)Z(g) \geqq 0$ and since $Z(d) > 0$ and $Z(x)$ does not vanish on G, we have $Z(x) > 0$ on G and hence $Y'(g) \geqq 0$. From (8) (a), using $b = g$, we have $Y'(x) > 0$ on G and hence $Y(x) > 0$ on D, thereby contradicting $Y(x) < 0$ on D. If $Z(g) = 0$, then $U(g) = Z'(g)Y(g) \geqq 0$ or $Z'(g) \geqq 0$, since $Y(g) < 0$. This, however, contradicts $Z'(g) < 0$ and completes the proof of our assertion that neither Y nor Z vanish on C.

From the non-vanishing of Y and Z on C follows that $Z'(c)$ is negative

negative. Let us show that $z'(x)$ is negative on C. Suppose that for a point $x=e$ of C we have $z'(e)=0$. Let $W(x)$ be the solution of the system $W''+p(x)W=0$, $W(e)=1$, $W'(e)=0$, and let $y=Wy^*$, $z=Wz^*$. Since $z'(e)=0$, we have

$$
(9)\qquad
\begin{aligned}
y^{*\prime}(x) &= [\int_e^x q(t)W(t)^2z^*(t)\,dt + y^{*\prime}(e)W(e)^2]/W(x)^2\\
z^{*\prime}(x) &= [\int_e^x r(t)W(t)^2y^*(t)\,dt]/W(x)^2.
\end{aligned}
$$

Since $y^*<0$ on E: $e \leqq x < \infty$ (y^* has the same sign as y) and since $z^*(e)=z/W>0$, we have that $z^*(x)$ is a positive increasing function on E. Hence $z^*(x) \geqq z^*(e) > K > 0$ and

$$y^*(x) \geqq y^*(e) + \int_e^x [K\int_e^s q(t)W(t)^2dt + y'(e)]ds/W(s)^2.$$

Applying Lemma II, we get $\lim_{x\to\infty} y^* = +\infty$. But since $y^*(e)=y(e)/W(e) < 0$ it follows that y^* and y must vanish on E, thereby contradicting our hypothesis that y does not vanish on D. Hence for every x on C, we have $z'(x)<0$. Furthermore, $y'(x)$ does not vanish on C. For, if $y'(h)=0$, then from $V(h)=-y(h)z'(h)\geqq 0$ follows that $z'(h)$ is positive or zero which contradicts $z'(c)<0$ on C. A similar argument shows that $Y'(x)$ cannot vanish on C. Since $Y'(c)=y'(c)$ and $Y'(d)>0$, it follows that $y'(c)$ is positive or zero on C. Since $y(x)$ is a negative increasing function and $z(x)$ is a positive decreasing function on C, it follows that $R(x)^2 = y(x)^2+z(x)^2$ is a decreasing function on C and is consequently less than a positive constant K. Now

$$
\begin{aligned}
\theta'(x) = -V(x)/R(x)^2 = -V(d)/R(x)^2 - [\int_d^x (qz^2-ry^2)\,dt]/R(x)^2\\
< -V(d)/K < -H < 0,
\end{aligned}
$$

where H is a constant. Hence $-\theta = -\theta(d) - \int_d^x \theta'(t)\,dt > -\theta(d) + H(x-d)$. Hence $\lim_{x\to\infty}\theta = -\infty$ and $y=R\cos\theta$, $z=R\sin\theta$ vanish infinitely often on D. This, however, contradicts the hypothesis that $y(x)$ and $z(x)$ each vanish only a finite number of times on D.

We omit the arguments for (B) since they parallel, step by step, the arguments for (A). In fact, if one wishes to do so he can bring this case under (A) by the use of the transformation $y=-z^*(x)$, $z=y^*$.

THEOREM VIII. *If there exists a positive constant J such that $|q|/(-p)^{1/2} \geqq J$, $|r|/(-p)^{1/2} \geqq J$ on X and if $(q+r)^2 - 4p^2 \leqq 0$ on X, then every solution $y(x)$, $z(x)$ of (3) either oscillates infinitely often on X or else the quantity $R^2 = y^2 + z^2$ decreases steadily approaching zero on some interval $G\colon g \leqq x < \infty$.*

Proof. Here again we confine our attention to the case $q > 0$ since the same arguments treat the case $q < 0$. If for any value of x, $x = d$, we have $V(d) \geqq 0$, then $V(x) > 0$ on $d < x < \infty$ and Theorem VII applies to prove that the solution oscillates infinitely often on X.

If there is a solution $y(x)$, $z(x)$ of (3) such that $V(x) < 0$ on X, we let d be chosen so that neither y nor z vanishes on $D\colon d \leqq x < \infty$. If no such d exists, then y and z each vanish infinitely often on X and the theorem is valid. We have

$$(R^2)' = 2(yy' + zz'), \qquad (R^2)'' = 2(yy'' + zz'' + y'^2 + z'^2).$$

If we multiply the equations of (3) by y and z, respectively, and add the resulting equations, we get

$$\tfrac{1}{2}(R^2)'' = y'^2 + z'^2 + Q,$$

where $Q = -py^2 + (q+r)yz - pz^2$ is a positive semi-definite quadratic form. Since $V(x)$ does not vanish on X, y' and z' cannot vanish simultaneously and $(R^2)''$ does not vanish on X. Hence $(R^2)'$ can vanish at most once on X. Choose a point $x = g$ on D so that $(R^2)'$ does not vanish on $G\colon g \leqq x < \infty$. Either $yz > 0$ on G or $yz < 0$ on that interval and in either case y and z do not vanish on G. An examination of (3) along with the inequalities $p < 0$, $q > 0$, $r < 0$ reveals that either y'' or z'' does not vanish on G. For the sake of definiteness, suppose that y'' does not vanish on G, then y' can vanish at most once on that interval. Let h be chosen greater than g so that y' does not vanish on $H\colon h \leqq x < \infty$. We show that neither y^2 nor z^2 can be bounded from zero on H. It follows from our hypotheses on $|q|/(-p)^{1/2}$ and $|r|/(-p)^{1/2}$ that $\int_h^x q(t)\,dt$ and $-\int_h^x r(t)\,dt$ become infinite with x and hence if either y^2 or z^2 were bounded from zero, the formula $V(x) = V(h) + \int_h^x$ [illegible] would give $\lim_{x\to\infty} V(x) = +\infty$, which contradicts $V(x) < 0$ on X. Since $y(h) \neq 0$, it follows that $yy' < 0$ on H and hence y^2 approaches zero steadily on H. Since z^2 cannot be bounded

is negative on this interval and since every neighborhood of infinity contains points where R^2 is arbitrarily small, is follows that R^2 approaches zero steadily on G.

COROLLARY. *If we omit the hypothesis* $(q+r)^2 - 4p^2 \leqq 0$ *from Theorem VIII, then we conclude that every solution of* (3) *either oscillates infinitely often on* X *or else* R^2 *takes on values arbitrarily close to zero in every neighborhood of infinity, i. e.,* $\varliminf_{x\to\infty} R^2 \doteq 0$, *where the underline indicates the lower limit.**

4°. *Case where the independent variable ranges over the entire real axis.* Consider the differential system (3) defined on the entire real axis $X: -\infty < x < +\infty$. Let qr be negative on X and $q(x)$, $r(x)$, and $p'(x)$ be continuous on this interval.

THEOREM IX. *If* $p(x) \equiv 0$ *on* C: $c \geqq x > -\infty$, *and the integrals* $\int_c^x tq(t)\,dt$, $\int_c^x tr(t)\,dt$ *become infinite as* x *becomes infinite on* C, *then every solution of* (3) *such that* $V(c)q(c) \leqq 0$ *will oscillate infinitely often on* C.

Proof. Carry the interval C into $T: -c \leqq t < +\infty$ by the change of independent variable $x = -t$. System (3) becomes

$$(10) \qquad d^2Y/dt^2 + P(t)Y = Q(t)Z, \qquad d^2Z/dt^2 + P(t)Z = S(t)Y,$$

where $P(t) = p(-t)$, $Q(t) = q(-t)$, $S(t) = r(-t)$, $Y(t) = y(-t)$, $Z(t) = z(-t)$. Let $U(t) = Y'(t)Z(t) - Z'(t)Y(t)$, then $V(x) = -U(t)$ and $U(t) = U(-c) + \int_{-c}^{t} [QZ^2 - SY^2]\,ds = -V(c) + \int_{-c}^{t} [QZ^2 - SY^2]\,ds$. Since $-V(c)q(c) \geqq 0$, it follows that $U(t)$ is a positive increasing function on $-c < t < \infty$ or a negative decreasing function on this interval according as $q(x)$ is positive or negative on C. Also, from the hypotheses on the integrals of $tq(t)$ and $tr(t)$, we have that $\int_{-c}^{t} sQ(s)\,ds$, $\int_{-c}^{t} sS(s)\,ds$ become infinite with t. We may now apply Theorem V to show that the solutions of (10) under investigation oscillate infinitely often on T and from this we immediately conclude that the corresponding solutions of (3) oscillate infinitely often on C.

---

* The corollary follows when we note that the only use made of the hypothesis $(q+r)^2 - 4p^2 \leqq 0$ in proving Theorem VIII was to show that $(R^2)'$ could not vanish on G.

The same transformation that was used in proving Theorem IX may be used along with Theorem VII to prove

THEOREM X. *If on C: $c \geqq x > -\infty$, we have $p(x) < 0$, $p'(x) \geqq 0$ and the quantities $q(x)/[-p(x)]^{1/2}$, $r(x)/[-p(x)]^{1/2}$ are bounded from zero, then every solution of* (3) *such that $q(c)V(c) \leqq 0$ oscillates infinitely often on C.*

If to the hypotheses of Theorems IX and X we add that $\int_c^\infty q(t)\,dt$, $\int_c^\infty r(t)\,dt$ become infinite as x becomes negatively infinite and $(r+q)^2 - 4p^2 \leqq 0$ on C, respectively, then Theorems VI and VIII are valid on C.

The results of this part of the paper are summarized in the following theorems:

THEOREM XI. *If in system* (3) *we have p, q, and r continuous on X: $-\infty < x < +\infty$, $qr < 0$ on X, and either:* (A), *$p(x) \equiv 0$ for $|x| > c$, where c is any fixed real number, and $\int_c^\infty tq(t)\,dt$, $\int_c^{-\infty} tq(t)\,dt$, $\int_c^\infty tr(t)\,dt$, $\int_c^{-\infty} tr(t)\,dt$ become infinite with x; or* (B), *$p(x) < 0$ for $|x| > c$, $p'(x) \leqq 0$ for $x \geqq c$, $p'(x) \geqq 0$ for $x \leqq -c$, $p'(x)$ continuous for $|x| > c$, and the quantities $q(t)/[-p(t)]^{1/2}$, $r(t)/[-p(t)]^{1/2}$ bounded from zero for $|x| > c$, then every solution of* (3) *oscillates infinitely often on X. Furthermore, if $x = a$ is any point of X and $V(a)q(a) \geqq 0$, the solution oscillates infinitely often on the positive real axis, while if $V(a)q(a) \leqq 0$, it oscillates infinitely often on the negative real axis. In particular, if for any solution $y(x)$, $z(x)$ of* (3) *we have $V(x) = 0$ at a point of X, then the solution oscillates infinitely often on both the positive and negative real axes.*

THEOREM XII. *If in addition to the hypotheses of Theorem XI we have in case* (A) *$\int_c^\infty q(t)\,dt$, $\int_c^{-\infty} q(t)\,dt$, $\int_c^\infty r(t)\,dt$, $\int_c^{-\infty} r(t)\,dt$ become infinite with x and in case* (B) *$(q+r)^2 - 4p^2 \leqq 0$, then every solution of* (3) *such that $V(a)q(a) \geqq 0$ oscillates infinitely often in a neighborhood of $+\infty$ and either oscillates infinitely often in a neighborhood of $-\infty$ or there exists a neighborhood of $-\infty$ in which $R^2 = y^2 + z^2$ approaches zero steadily and every solution of* (3) *such that $V(a)q(a) \leqq 0$ oscillates infinitely often in a neighborhood of $-\infty$ and either oscillates infinitely often in a neighborhood of $+\infty$*

The geometrical characterization of the integral curve of (3) in the yz-plane is one of the following:

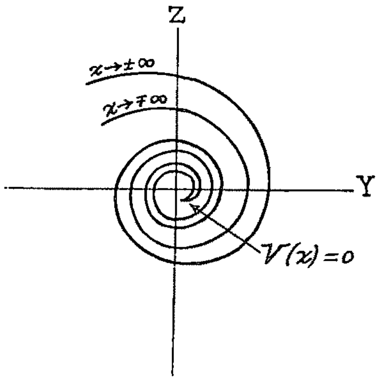


FIGURE I.

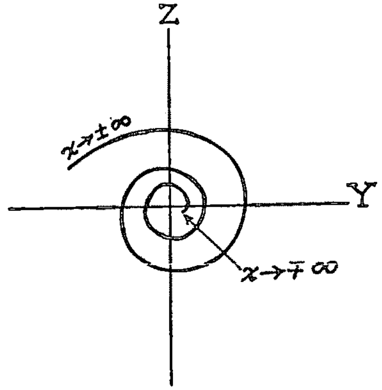


FIGURE II.

III. SYSTEM (3) WITH $q(x)r(x) > 0$.

Consider system (3) where $p(x)$, $q(x)$ and $r(x)$ are of class C' on X: $a \leqq x \leqq b$ and $q(x)r(x) > 0$ on this interval.

Let $w_1(x)$ and $w_2(x)$ be linearly independent solutions of (6) and let $W[w_1, w_2] = w_1(x)w'_2(x) - w_2(x)w'_1(x)$ be the wronskian of these solutions. One easily shows that $W[w_1, w_2]$ is constant on X and is different from zero. If $G(x, t)$ is defined by

$$G(x, t) = [w_1(t)w_2(x) - w_1(x)w_2(t)]/W[w_1, w_2],$$

then the general solution of (3) is given by

$$\begin{aligned} y(x) &= c_1w_1(x) + c_2w_2(x) + \int_a^x G(x,t)q(t)z(t)\,dt \\ z(x) &= c_3w_1(x) + c_4w_2(x) + \int_a^x G(x,t)r(t)y(t)\,dt, \end{aligned} \tag{11}$$

where c_1, c_2, c_3, c_4 are constants.

THEOREM XIII. *If the interval X is such that there exists a non-vanishing solution of (6) on it, then w_1 and w_2 can be so chosen that $W[w_1, w_2] \equiv 1$, $G(x, t) > 0$ on $a \leqq t < x$, and $G(x,t) < 0$ on $x < t \leqq b$.*

Proof. If $w_0(x)$ is the non-vanishing solution of (6) and we let $w_1(x) \equiv | w_0(x) | / | w_0(a) |$ while $w_2(x)$ is the solution of (6) such that

$w_2(a) = 0$, $w'_2(a) = 1$, then $G(x, t)$ and $W[w_1, w_2]$ are readily seen to have the properties that are desired for Theorem XIII.

Let (y_1, z_1), (y_2, z_2), (y_3, z_3), (y_4, z_4) be linearly independent solutions of (3). It follows by well-known considerations that these four solutions form a fundamental system of solutions of (3) and that

$$W = \begin{vmatrix} y_1 & y_2 & y_3 & y_4 \\ y_1' & y_2' & y_3' & y_4' \\ z_1 & z_2 & z_3 & z_4 \\ z_1' & z_2' & z_3' & z_4' \end{vmatrix} \equiv \text{a constant} \neq 0.$$

Let u_i and v_i, $(i = 1, 2, 3, 4)$, be the cofactors in W of y_i' and z_i', respectively. Calculation of u_i', v_i', u_i'', v_i'' shows that (u_i, v_i), $(i = 1, 2, 3, 4)$, is a solution of

$$(12) \qquad u'' + pu = rv, \qquad v'' + pv = qu.$$

If W^* denotes the determinant obtained from W by changing y_i, y_i', z_i, z_i', respectively, to u_i, u_i', v_i, v_i', $(i = 1, 2, 3, 4)$, then * $W^* \equiv W^3$. W^* is therefore different from zero and (u_i, v_i), $(i = 1, 2, 3, 4)$, form a fundamental system of solutions of (12). If in W^* we let Y_i be the cofactor of u_i' and Z_i the cofactor of v_i', one easily verifies † that $y_i = -WY_i/W^* = -Y_i/W^2$, $z_i = -WZ_i/W^* = -Z_i/W^2$.

Let

$$K_{11}(x, t) = \sum_1^4 y_i(x)u_i(t)/W, \qquad K_{12}(x, t) = \sum_1^4 y_i(x)v_i(t)/W,$$

$$K_{21}(x, t) = \sum_1^4 z_i(x)u_i(t)/W, \qquad K_{22}(x, t) = \sum_1^4 z_i(x)v_i(t)/W.$$

The K's are readily seen to have the following properties:

$$\begin{array}{llll} K_{11}(x,x) = 0, & K_{12}(x,x) = 0, & K_{21}(x,x) = 0, & K_{22}(x,x) = 0 \\ (\partial/\partial x)K_{11}(x,x) = 1, & (\partial/\partial x)K_{12}(x,x) = 0, & (\partial/\partial x)K_{21}(x,x) = 0, & (\partial/\partial x)K_{22}(x,x) = 1 \\ (\partial/\partial t)K_{11}(x,x) = -1, & (\partial/\partial t)K_{12}(x,x) = 0, & (\partial/\partial t)K_{21}(x,x) = 0, & (\partial/\partial t)K_{22}(x,x) = - \end{array}$$

---

* Cf. Bôcher, *Introduction to Higher Algebra*, New York (1919), p. 33.

† This is done by solving the following sets of algebraic equations:

$$\sum_1^4 u_iY_i = 0, \quad \sum_1^4 u_i'Y_i = -W^*, \quad \sum_1^4 v_iY_i = 0, \quad \sum_1^4 v_i'Y_i = 0,$$

$$\sum_1^4 u_iy_i = 0, \quad \sum_1^4 u_i'y_i = -W, \quad \sum_1^4 v_iy_i = 0, \quad \sum_1^4 v_i'y_i = 0,$$

For each fixed x on X, (K_{11}, K_{12}) and (K_{21}, K_{22}) are solutions of (12). For each fixed t on X, (K_{11}, K_{21}) and (K_{12}, K_{22}) are solutions of (3).

THEOREM XIV. *The functions K_{11}, K_{12}, K_{21}, K_{22} are uniquely determined for system (3), i. e., they are independent of the particular fundamental system of solutions used in their construction.*

Proof. Let $K^*_{ij}(x, t)$, $(i, j = 1, 2)$, be formed for a different fundamental system of solutions of (3) and suppose that $K^*(c, d) \neq K(c, d)$ for some values of i and j and for some c and d such that $a \leqq c$, $d \leqq b$. $K_{1j}(x, d)$, $K_{2j}(x, d)$ and $K^*_{1j}(x, d)$, $K^*_{2j}(x, d)$ are solutions of system (3) and both have the same initial values at $x = d$. Hence by the uniqueness part of the fundamental existence theorem, $K_{ij}(x, d) = K^*_{ij}(x, d)$ on $a \leqq x \leqq b$, thereby contradicting $K_{ij}(c, d) \neq K^*_{ij}(c, d)$.

THEOREM XV. *The functions K_{ij} have the following properties:*

$$K_{11}(x, t) = -K_{22}(t, x), \quad K_{21}(x, t) = -K_{21}(t, x),$$
$$K_{12}(x, t) = -K_{12}(t, x), \quad K_{22}(x, t) = -K_{11}(t, x).$$

Proof. Following Theorem XIV we can and will form the K's from a fundamental system of solutions that has the property of making $W \equiv 1$. $y_i = v_i$, $z_i = u_i$ is a solution of (3), since u_i, v_i is a solution of (12) and $W = W^* = 1$. Hence the cofactors of v_i' and u_i' in W are respectively $-z_i$ and $-y_i$. Substituting we obtain

$$K_{11}(x, t) = \sum_1^4 y_i(x) u_i(t) = -\sum_1^4 v_i(x) z_i(t) = -K_{22}(t, x),$$

$$K_{12}(x, t) = \sum_1^4 y_i(x) v_i(t) = -\sum_1^4 v_i(x) y_i(t) = -K_{12}(t, x),$$

$$K_{21}(x, t) = \sum_1^4 z_i(x) u_i(t) = -\sum_1^4 u_i(x) z_i(t) = -K_{21}(t, x),$$

$$K_{22}(x, t) = \sum_1^4 z_i(x) v_i(t) = -\sum_1^4 u_i(x) y_i(t) = -K_{11}(t, x).$$

Theorem XV is important in that it gives the values of the K's at every point of $x < t \leqq b$ in terms of the values of these functions on $a \leqq t < x$.

The general solution of

$$(13) \qquad y'' + py - qz = h_1(x) z, \qquad z'' + pz - ry = h_2(x) y,$$

where h_1 and h_2 are continuous real functions of class C' on X and $(q + h_1)(r + h_2) \neq 0$ on X, is given by the pair of integral equations

$$
(14)\qquad
\begin{aligned}
y(x) &= \sum_1^4 c_i y_i(x) + \int_a^x [K_{11}(x,t)h_1(t)z(t) + K_{12}(x,t)h_2(t)y(t)]\,dt \\
z(x) &= \sum_1^4 c_i z_i(x) + \int_a^x [K_{21}(x,t)h_1(t)z(t) + K_{22}(x,t)h_2(t)y(t)]\,dt,
\end{aligned}
$$

where c_1, c_2, c_3, c_4 are constants. The equivalence of systems (13) and (14) may be verified by substitution.

*Definition. The interval X is said to be regular** *of the first kind for* (3) *if $K_{ij}(x, t) > 0$, $(i, j = 1, 2)$, for each x on X^*: $a < x \leqq b$ and every t on $a \leqq t < x$.*

Definition. The interval X is regular of the second kind for (3) *if $K_{11} > 0$, $K_{22} > 0$, $K_{12} < 0$, $K_{21} < 0$, for each x on X^* and for every t on $a \leqq t < x$.*

THEOREM XVI. *If $r > 0$, $q > 0$, on X and there exists a solution of* (6) *that does not vanish on X, then X is regular of the first kind for* (3).

Proof. Let $G(x, t)$ be defined as indicated under Theorem XIII and assume that for some value of x, $x = c$, on X^* and for some t, $t = d$, on $a \leqq t < c$, at least one K_{ij} is less than or equal to zero. Since $K_{11}(c, t)$, $K_{12}(c, t)$ is a solution of (12), we have

$$
(15)\qquad
\begin{aligned}
&\text{(a)}\quad K_{11}(c,t) = -w_2(t) + \int_c^t G(t,s)r(s)K_{12}(c,s)\,ds \\
&\text{(b)}\quad K_{12}(c,t) = \int_c^t G(t,s)q(s)K_{11}(c,s)\,ds.
\end{aligned}
$$

Since $w_2(t) < 0$ for $t < c$ and $(\partial/\partial t)K_{11}(c, c) = -1$, $K_{11}(c, c) = 0$, it follows that for a sufficiently small positive ϵ, $K_{11}(c, t) > 0$ on $c - \epsilon \leqq t < c$. An examination of equation (15) (b) shows that the integrand is negative on the interval $c - \epsilon \leqq t < c$ for $s > t$ and since the upper limit of integration is smaller than the lower, it follows that $K_{12}(c, t)$ is positive on this interval. If either $K_{11}(c, t)$ or $K_{12}(c, t)$ vanishes at $t = d$, we let $t = e$ be the greatest zero of the product of these two functions on $d \leqq t < c$. If $K_{11}(c, e) = 0$, then equation (15) (a) yields

$$
(16)\qquad 0 = -w_2(e) + \int_c^e G(e,s)r(s)K_{12}(c,s)\,ds.
$$

Both terms on the right hand side of equation (16) are positive and hence this equation is impossible. Similar reasoning when applied to equation [illegible]

* Cf. Birkhoff, *loc. cit.*, page 113, and Reynolds, *loc. cit.*, page 223.

Since $K_{21}(c,t)$, $K_{22}(c,t)$ is a solution of (12), we have

$$K_{21}(c,t) = \int_c^t G(t,s)r(s)K_{22}(c,s)\,ds \tag{17}$$
$$K_{22}(c,t) = -w_2(t) + \int_c^t G(t,s)q(s)K_{21}(c,s)\,ds$$

and an application of reasoning that is entirely analogous to that used in treating K_{11} and K_{12} shows that neither $K_{21}(c,d)$ nor $K_{22}(c,d)$ can be zero. Hence none of the K's can vanish at (c,d) and it follows from this contradiction that Theorem XVI must hold.

Obvious modifications of the proof of Theorem XVI establish:

THEOREM XVII. *If $r < 0$, $q < 0$ on X and if there exists a solution of* (6) *that does not vanish on X, then X is regular of the second kind for* (3).

We note as an obvious consequence of the foregoing work that any regular interval for (3) is also regular and of the same kind for (12) and conversely, any regular interval for (12) is also regular and of the same kind for (3).

In the presence of the requirement $rq > 0$ it is true that every finite interval X can be broken into a finite number of regular intervals for (3) all of which are of the first kind or else they all are of the second kind. One merely notes that a solution of (6) has only a finite number of zeros on any interval of finite length. In particular, if $p(x) < 0$ equation (6) has solutions that do not vanish on any sub-interval of the real axis and in this case the entire real axis is regular of one kind for (3) and (12).

THEOREM XVIII. *If X is a regular interval of the first (second) kind for system* (3) *and (Y, Z) and (y, z), respectively, are non-identically vanishing solutions of* (3) *and* (13) *such that $y(c) = Y(c)$, $y'(c) = Y'(c)$, $z(c) = Z(c)$, $z'(c) = Z'(c)$, $a \leqq c \leqq b$, and if $h_1(x) \equiv 0$ on X, then the first zero* of $y(x)$ on $c < x \leqq b$ follows (precedes) or precedes (follows)*

* The proof that the zeros of y, z, Y, and Z on X cannot have a finite limit point is as follows:

Assume that $x = c$ is a limit point of the zeros of one of these quantities, for example $y(x)$. Rolle's theorem and continuity considerations show that $x = c$ is also a limit point of zeros of $y'(x)$, $y''(x)$, and $y'''(x)$ and furthermore, $y(c) = y'(c) = y''(c) = y'''(c) = 0$. Now $z(c) = [y''(c) + p(c)y(c)]/q(c) = 0$, $z'(c) = [y'''(c) + p'(c)y(c) + p(c)y'(c) - q'(c)z(c)]/q(c) = 0$. Hence $y(c) = y'(c) = z(c) = z'(c) = 0$ and by the fundamental existence theorem, $y(x) \equiv z(x) \equiv 0$ on X. This contradicts $y(x)$ and $z(x)$ not identically zero on X. Similar arguments when applied to the proper equations of (3) or (13) treat the cases $z(x)$, $Y(x)$, $Z(x)$.

the first zero of $Y(x)$ on this interval according as $h_2(x) \geqq 0$ on $c < x \leqq b$ or $h_2(x) \leqq 0$ on this interval, in any case the equality sign holding on at most a null set of points.

Proof. Let a in formula (14) be replaced by c and note that $\sum_1^4 c_i y_i(x)$, $\sum_1^4 c_i z_i(x)$ is a solution of (3) that has the same initial conditions at $x = c$ as $Y(x)$, $Z(c)$. Using the uniqueness part of the fundamental existence theorem for (3), system (14) becomes

$$(18) \qquad \begin{aligned} &\text{(a)} \quad y(x) = Y(x) + \int_c^x K_{12}(x, t) h_2(t) y(t)\, dt \\ &\text{(b)} \quad z(x) = Z(x) + \int_c^x K_{22}(x, t) h_2(t) y(t)\, dt. \end{aligned}$$

It follows readily by differentiation that the lowest derivatives of y and Y that do not vanish at $x = c$ are equal at this point. Hence if $x = d$ is a point such that $d > c$ and neither y nor Y vanish on $c < x \leqq d$, then on this interval $Y(x)y(x)$ is positive.

If $h_2(x) > 0$ almost everywhere on X and X is regular of the first kind, we have that $\int_c^x K_{12}(x) h_2(t) y(t)\, dt$ is different from zero and has the same sign as $Y(x)$ for every x on $c < x \leqq b$ that is less than or equal to the smallest value of x on this interval for which yY vanishes. Hence the product yY must vanish for a smaller value of x on $c < x \leqq b$ than the smallest value on this interval for which $y(x)$ vanishes and our theorem follows for this case.

If X is regular of the second kind, we transpose the integral term of (18) (a) to the other side and repeat the above argument with respect to $Y(x)$ to get the desired conclusion. If $h_2(x) \leqq 0$ on X, a repetition of the foregoing argument with obvious modifications treats all of the cases that arise here.

If $h_2(x) \equiv 0$ on X and (y, z), (Y, Z) are defined as in Theorem XVIII, equations (14) become

$$(19) \qquad \begin{aligned} y(x) &= Y(x) + \int_c^x K_{11}(x, t) h_1(t) z(t)\, dt \\ z(x) &= Z(x) + \int_c^x K_{21}(x, t) h_1(t) z(t)\, dt \end{aligned}$$

and reasoning entirely analogous to that used in proving Theorem XVIII

THEOREM XIX. *If X is a regular interval of the first (second) kind for (3) and $h_2(x) \equiv 0$ on X, then the first zero of $z(x)$ on $c < x \leqq b$ follows (precedes) or precedes (follows) the first zero of $Z(x)$ on that interval according as as $h_1(x) > 0$ almost everywhere on $c < x \leqq b$ or $h_1(x) < 0$ almost everywhere on that interval.*

Using the properties of $K_{ij}(x, t)$, $(i, j = 1, 2)$, that were established in Theorems XIV and XV, we establish the following theorems on $a \leqq x < c$:

THEOREM XX. *Under the hypotheses of Theorem XIX, if X is a regular interval of the first (second) kind for (3), the last zero of $y(x)$ on $a \leqq x < c$ precedes (follows) or follows (precedes) the last zero of $Y(x)$ on this interval according as $h_2(x) > 0$ almost everywhere on $a \leqq x < c$ or $h_2(x) < 0$ almost everywhere on this interval.*

THEOREM XXI. *Under the hypotheses of Theorem XIX, if X is a regular interval of the first (second) kind for (3), the last zero of $z(x)$ on $a \leqq x < c$ precedes (follows) or follows (precedes) the last zero of $Z(x)$ on that interval according as $h_1(x) > 0$ almost everywhere on $a \leqq x < c$ or $h_1(x) < 0$ almost everywhere on that interval.*

THEOREM XXII. *Let (y, z) and (Y, Z) be defined as in Theorem XVIII and $q(c)[q(c) + h_1(c)] > 0$, $r(c)[r(c) + h_2(c)] > 0$. If X is regular of the first (second) kind for (3) and the first non-vanishing derivative of $y(x)$ at $x = c$ has the same (opposite) sign as the first derivative of $z(x)$ that does not vanish at $x = c$, then the first zero of yz on $c < x \leqq b$ follows (precedes) or precedes (follows) the first zero of YZ on this interval according as $h_1(x) > 0$, $h_2(x) > 0$ almost everywhere on $c < x \leqq b$ or $h_1(x) < 0$, $h_2(x) < 0$ almost everywhere on that interval.*

Proof. Following the reasoning used under Theorem XVIII we may write the lower limits of integrations in system (14) as c and replace $\sum_1^4 c_i y_i(x)$ by $Y(x)$ and $\sum_1^4 c_i z_i(x)$ by $Z(x)$ in this system. Calculation of the derivatives of Y, Z, y, and z together with the hypotheses on $q(c)[q(c) + h_1(c)]$ and $r(c)[r(c) + h_2(c)]$ shows that if $x = d$ is any point such that $yzYZ \neq 0$ on $c < x \leqq d$, then $y(d)Y(d) > 0$, $z(d)Z(d) > 0$. An examination of the right hand side of (14) shows that if X is regular of the first kind, and the lowest non-vanishing derivatives of $y(x)$ and $z(x)$ at $x = c$ are of the same sign, the integrals will have the same signs (opposite

signs) as $Y(x)$ and $Z(x)$ if $h_1 > 0$, $h_2 > 0$ ($h_1 < 0$, $h_2 < 0$) almost everywhere on X. Application of an argument of the type of that was used in proving Theorems XVIII and XIX shows that the first zero of yz on $c < x \leqq b$ must precede the first zero of YZ on this interval.

An entirely similar argument demonstrates the theorem for the cases where X is regular of the second kind.

An argument that follows exactly the same lines as that used in proving Theorem XXII establishes the following theorem:

THEOREM XIII. *Let $q(c)[q(c) + h_1(c)]$ and $r(c)[r(c) + h_2(c)]$ be positive and let y, z, Y, Z, be defined as in Theorem XVIII. Let $x = e$ be the first zero of yz on $c < x \leqq b$ and let $x = f$ be the first zero of YZ on this interval. Let X be regular of the first (second) kind for (3) and let the lowest non-vanishing derivatives of y and z at $x = c$ have opposite (the same) signs. Then if $\left\{\begin{array}{l} z(e) = 0 \\ y(e) = 0 \end{array}\right\}$, we have $\left\{\begin{array}{l} f < e \ (f > e) \\ f > e \ (f < e) \end{array}\right\}$ or $\left\{\begin{array}{l} f > e \ (f < e) \\ f < e \ (f > e) \end{array}\right\}$ according as $h_1 > 0$, $h_2 < 0$ almost everywhere on $c < x \leqq b$ or $h_1 < 0$, $h_2 > 0$ on that interval.*

It is evident that Theorems XIV and XV may be used to establish theorems on $a \leqq x < c$ that are the exact analogues of Theorems XXII and XXIII.

Definition. By the $\left\{\begin{array}{l} \textit{forward} \\ \textit{backward} \end{array}\right\}$ interval of oscillation of (3) at $x = c$ we mean the subinterval $\left\{\begin{array}{l} c \leqq x \leqq d \\ e \leqq x \leqq c \end{array}\right\}$ of X, if such exists, such that $\left\{\begin{array}{l} d \\ e \end{array}\right\}$ is the $\left\{\begin{array}{l} \textit{smallest} \\ \textit{largest} \end{array}\right\}$ number having the property that every solution of (3) whose y or z vanishes at $x = c$ also has a zero of $y(x)z(x)$ on $\left\{\begin{array}{l} c < x \leqq d \\ e \leqq x < c \end{array}\right\}$.*

THEOREM XXIV. *If $q(x)r(x) > 0$ on X, the backward and forward intervals of oscillation for (3), if they exist, are greater in length than the corresponding intervals† for (6).*

---

* Cf. Birkhoff, *loc. cit.*, page 119.

† These intervals for (6) are determined by the zeros that immediately precede

Proof. Let $w(x)$ and $W(x)$ be the solutions of (6) having the values $w(c)=0$, $w'(c)=1$, $W(c)=1$, $W'(c)=0$, where $x=c$ is any point of X. and form $G(x,t)$ from these solutions. $y(x)=\int_c^x G(x,t)q(t)z(t)\,dt+w(x)$, $z(x)=\int_c^x G(x,t)r(t)y(t)\,dt$ represents the solution of (3) such that $y(c)=z(c)=z'(c)=0$, $y'(c)=1$. If $x=d$ is the first zero of $w(x)$ such that $d>c$, then $G(d,t)$ is positive on $c<x<d$ and vanishes at $x=c$ and $x=d$. An examination of the above equations for $y(x)$ and $z(x)$ shows that neither y nor z can vanish on $c<x\leqq d$. Similar considerations treat the backward intervals of oscillation.

COROLLARY. *If $qr>0$ and (6) has a non-oscillatory solution on any interval X, there exist solutions of (3) such that either y or z does not vanish on X.*

IV. APPLICATION TO EQUATIONS (1) AND (2).

We indicate briefly how the foregoing treatment applies to the general fourth order self-adjoint equation (1) and to the second order equation (2).

For equation (1) we have $p(x)=L(x)/2K(x)$, $q(x)=1/K(x)$, $r(x)=L''(x)/2+L(x)^2/4K(x)-G(x)$. The conditions that we have imposed on p, q, and r in the various theorems are conditions on the above combinations of the coefficients of (1). The conclusions of these theorems apply to $u(x)$, the solution of (1), and the combination $Ku''+Lu/2$. In the special but important case $L(x)=0$, we have $p(x)=0$, $r(x)=-G(x)$ and the conditions of the theorems become very simple.

Equation (2) is equivalent to system (3) where $q(x)=-r(x)$. This property simplifies several of the hypotheses of the theorems of the paper. In particular, the condition $(q+r)^2-4p^2\leqq 0$ is automatically satisfied. Let $u(x)=u_1(x)+iu_2(x)$ in equation (2), then

$$
\begin{aligned}
u_1 \exp\left[\tfrac{1}{2}\int_a^x p_1\,dt\right] &= y\cos\int_a^x p_2\,dt/2+z\sin\int_a^x p_2\,dt/2\\
u_2 \exp\left[\tfrac{1}{2}\int_a^x p_1\,dt\right] &= -y\sin\int_a^x p_2\,dt/2+z\cos\int_a^x p_2\,dt/2
\end{aligned}
\tag{20}
$$

and the determinant of coefficients of y and z is identically unity.

The following theorem concerning equations of type (20) forms a connecting link between the properties of y and z and those of u_1 and u_2.

THEOREM XXV. *If $y(x)$, $z(x)$, $a(x)$, and $b(x)$ are of class C' on X and $(a'b-ab')(y'z-yz')\geqq 0$, $V(x)=y'z-z'y\neq 0$, $a\neq 0$, $b\neq 0$, on X, then*

1). *The zeros of $y(x)$ and $z(x)$ separate each other on X.*

2). *If $u=ay-bz$, $v=by+az$, there is at least one zero each of u and v between each pair of zeros of either y or z on X.*

3). *The zeros of u and v separate each other on X.*

4). *There is only one zero each of u and v between each pair of consecutive zeros of y and between each pair of consecutive zeros of z.*

Proof. 1). Let C and D denote two zeros of y on X and assume that z does not vanish between C and D. $z(x)$ cannot vanish at either C or D since this would cause V to vanish at such a point. Consider the continuous function y/z on the closed interval bounded by C and D. The derivative $[y/z]'=V(x)/z^2$ does not vanish on this interval and hence by Rolle's theorem y/z cannot vanish at both C and D. This contradiction shows that z vanishes at least once between each pair of zeros of y. Similar considerations of z/y shows that y must vanish at least once between each pair of zeros of z.

2). $y(x)$ and $z(x)$ cannot have a finite limit point for their zeros on X, for continuity considerations together with conclusion 1) above would cause both $y(x)$ and $z(x)$ to vanish at such a point and hence $V(x)$ would necessarily vanish there. Let $x=c$ and $x=d$ be consecutive zeros of y on X and assume that $v(x)$ does not vanish on $c<x<d$. An examination of the equations connecting y, z, u, and v, together with the restriction on $V(x)$, shows that no two of these quantities can vanish simultaneously on X. Since $z(x)$ vanishes exactly once on $c<x<d$, we have from $a(x)\neq 0$, $v(c)=a(c)z(c)$, $v(d)=a(d)z(d)$, that $v(c)v(d)<0$ and hence $v(x)$ must vanish at least once on $c<x<d$. Similar arguments treat all of the other cases.

3). Since u and v do not vanish simultaneously, we establish 3) by considerations similar to those used in proving 1). We note that $(u/v)'=[(a'b-ab')(y^2+z^2)+(a^2+b^2)(y'z-yz')]/v^2$, when $v(x)\neq 0$, and this quantity does not vanish on X.

4). In the presence of conclusion 3), we may use the same argument that was used under 2) and show that u and v cannot have finite limit points [illegible]

and $z(x)$, we obtain $y(x) = (au + bv)/(a^2 + b^2)$, $z(x) = (av - bu)/(a^2 + b^2)$. The argument that was used under 2) to show that v vanishes at least once between each pair of consecutive zeros of $y(x)$ can now be applied to show that $y(x)$ and $z(x)$ each vanish at least once between each pair of consecutive zeros of $u(x)$ and between each pair of consecutive zeros of $v(x)$. This together with 2) establishes part 4).

In conclusion, it is worthwhile to point out that the foregoing treatment does not touch the cases where $q(x)r(x)$ changes signs on the interval under consideration. The methods of the present paper do not seem to lend themselves readily to a treatment of these cases except in so far as it is possible to break up the given interval into sub-intervals upon each of which either $q(x)r(x) \leqq 0$ or $q(x)r(x) \geqq 0$. Other important cases omitted from the present paper are those that arise when $p(x)$ neither vanishes identically nor remains negative on interval considered. All of these cases are important and it would be of interest to develop the properties of the solutions of such systems.

HARVARD UNIVERSITY AND
UNIVERSITY OF CALIFORNIA AT LOS ANGELES.

Mapping of a General Type of Three Dimensional Region on a Sphere.*

By John J. Gergen.†

Part I. Mapping of the Interior.

1. *Introduction.* We shall be concerned in this paper with a three dimensional region D having the following properties:

1. It is bounded;
2. It is normal,‡ that is, a solution of the classical Dirichlet problem exists for D for any assigned continuous boundary values. In particular, if P_0 be any point of D, then the Green's function $g(D \mid P_0, P)$ with pole at P_0, exists for D;
3. There is a point O in D such that $\nabla\, g(D \mid O, P)$ vanishes at no point P in D.

We shall prove that this region D can be mapped upon the interior D' of a sphere in such a way that:

1. The images in D' of points on a level surface $g(D \mid O, P) =$ const. lie on a sphere concentric with D';

---

* Preliminary report presented to the American Mathematical Society, February, 1928, under the title of "On Accessible Points on the Boundary of a Three Dimensional Region." This problem was suggested by Professor G. C. Evans, and the author is particularly indebted to Professor Evans for suggesting—on the basis of physical intuition of the flow of electricity—a type of simply connected region for which $|\nabla g|$ vanishes at an interior point, and for the proof of Lemma 7, second paper, which takes the place of two lemmas given by Professor Osgood for the two dimensional situation.

† National Research Fellow.

‡ For a discussion of the methods of attack on the problem of Dirichlet, sufficient conditions for normality, and a complete set of references, see Kellogg, "Recent Progress with the Dirichlet Problem," *Bulletin of the American Mathematical Society*, Vol. 32 (1926), pp. 601-625. See also Kellogg, *Foundations of Potential Theory*, Berlin, [illegible] *Journal de Mathématiques Pures et Appliquées*, t. 7 (1928), pp. 93-111. In this memoir Bouligand discusses the problem of Dirichlet and the principle of Picard, employing in his analysis the trajectories orthogonal to the level surfaces of the "generalized" Green's [illegible]

2. The images of points on any trajectory orthogonal to these level surfaces lie on a radius of D';
3. The points in D correspond in a $1:1$ manner to those in D';
4. The mapping functions (of the coördinates) which carry D into D' are analytic in $D - O$ and continuous at O.

The importance of this type of mapping lies, of course, in its analogy to conformal mapping of plane regions upon a circle. This analogy rests on the fact that when a plane region S is mapped in a $1:1$ and conformal manner upon the interior S' of a circle, the level curves of the Green's function of S, with pole at the point corresponding to the center of S', are carried into circles concentric with S' and the trajectories orthogonal to these surfaces are carried into radii of S'.

In a second paper, to be published in a forthcoming issue of this Journal, we shall be interested in the correspondence brought about between the boundary points of D and those of D' when a correspondence of the above type is set up between the interior points. Our methods here are based essentially upon those employed by Osgood and Taylor in their proof of the fundamental theorem * that when a simply connected plane region S is mapped in a $1:1$ and conformal manner upon the interior of the circle S', then an accessible point A, on the boundary of S, corresponds to a single point on the boundary of S', in the sense that the image of the curve along which A is accessible is a curve with a single limit point in the boundary of S'.

1.1 In regard to the quantity $|\nabla g(D \mid O, P)|$ the situation is somewhat different in three dimensions than it is in two. In fact, if S is a simply connected plane region and $g(S \mid O, P)$ its Green's function, then $|\nabla g(S \mid O, P)| > 0$ at every point P of S, independently of where in S the pole O is chosen. On the other hand, if T is a simply connected, three dimensional, normal domain, it is possible that $|\nabla g(T \mid O, P)|$ vanishes at some point P for some position of the pole O. This is shown to be the case in the following example.

Let T be the open set of points bounded by the torus t, obtained by revolving the circle

$$(y-4)^2 + z^2 = 1, \qquad x = 0,$$

about the z-axis. Let V_h be the open set of points determined by the inequalities

* Osgood and Taylor, "Conformal Transformations on the Boundaries of Their Regions of Definition," *Transactions of the American Mathematical Society*, Vol. 14 (1913), pp. 277-298.

$$x^2 + y^2 < 16, \qquad |z| < h \leq 1,$$

where $h > 0$, and let $S_h = T + U_h$.

The regions S_h and T are normal by several known conditions for normality, and S_h is simply connected. On the other hand, we can prove that $|\nabla g(S_h \mid A, P)|$ vanishes at some point P_h for small values of h, A being the point $(0, -4, 0)$.

This phenomenon is not surprising when the function $g(S_h \mid A, P)$ is interpreted as the potential at P due to a charge of electricity at A, the boundary of S_h being kept at zero potential. For by decreasing h we increase the resistance of the plate $S_h - T$ to the flow of electricity through it, while we do not change directly the flow through T. Thus for sufficiently small values of h one might expect that in the region T the potential function $g(S_h \mid A, P)$ behaves sensibly like $g(T \mid A, P)$. The latter of course, has an equilibrium point on the positive y-axis, and hence it is conceivable that so also has $g(S_h \mid A, P)$. That something like this actually is true we prove below.

We observe that the function

$$G(S_h \mid P) = g(S_h \mid A, P) - g(T \mid A, P)$$

is harmonic in T, if properly defined at A, and is continuous in $T + t$. Accordingly, because of the regularity of t,*

$$G(S_h \mid P) = (1/4\pi) \iint_t G(S_h \mid Q) \left[\partial g(T \mid P, Q)/\partial n\right] d\sigma_Q,$$

where n is the interior normal to t, and Q is a point on the element of integration $d\sigma_Q$. This integral, of course, reduces to

$$G(S_h \mid P) = (1/4\pi) \iint_{\alpha(h)} g(S_h \mid A, Q) \left[\partial g(T \mid P, Q)/\partial n\right] d\sigma_Q, \tag{1}$$

where $\alpha(h)$ is the set of points on t determined by the inequalities

$$x^2 + y^2 \leq 16, \qquad |z| \leq h.$$

Now since $g(T \mid A, P)$ vanishes on t and is positive throughout T, we can find two points, P_1, P_2, on the segment of the positive y-axis interior to T, with $\overline{AP}_1 < \overline{AP}_2$, such that

$$g(T \mid A, P_1) - g(T \mid A, P_2) [illegible] \epsilon.$$

* Compare Goursat, *Cours d'Analyse*, Vol. 3 (1927), p. 523.

where ϵ is a sufficiently small positive quantity. At the point P_2 we have

$$g(S_h \mid A, P_2) > g(T \mid A, P_2),$$

since $G(S_h \mid P)$ is harmonic in T and non-negative on t, and at the point P_1 we have

$$g(S_h \mid A, P_1) - g(T \mid A, P_1) \leq J \text{ area } \alpha(h),$$

where J is a constant independent of h, since $\partial g(T \mid P_1, Q)/\partial n$ and $g(S_1 \mid A, Q)$ are continuous on t and the latter exceeds $g(S_h \mid A, Q)$ on $\alpha(h)$. Accordingly,

$$g(S_h \mid A, P_1) \leq J \text{ area } \alpha(h) - \epsilon + g(T \mid A, P_2).$$

Thus when h is small enough

$$g(S_h \mid A, P_1) < g(S_h \mid A, P_2). \tag{2}$$

From (2) and the fact that $\partial g(S_h \mid A, P)/\partial y < 0$ when P is on $\overline{AP_1}$, in the neighborhood of A, it follows that for some point P_h on $\overline{AP_2}$, $\partial g(S_h \mid A, P_h)/\partial y = 0$, when h is small enough. We can conclude from this the truth of the statement, for when P is on the y-axis we have, because of the position of the pole and the symmetry of S_h,

$$\mid \nabla g(S_h \mid A, P) \mid = \mid \partial g(S_h \mid A, P)/\partial y \mid .$$

2.1. *The Orthogonal Trajectories.* We proceed now to a discussion of the level surfaces, $g = \text{const.}$, and their orthogonal trajectories.

We choose the point O, the pole of $g(D \mid O, P)$ as the origin of a system of rectangular axes, x, y, z. The function g then takes the form

$$g(D \mid O, P) = g(P) = g(x, y, z) = 1/r + v(x, y, z),$$

where

$$r = (x^2 + y^2 + z^2)^{1/2}$$

and the function v is harmonic in D and continuous in the closed set $D + d$, d being the boundary of D.

Since g is harmonic and therefore analytic* in $D - O$, and since we

---

* The word "analytic" is used in the ordinary sense. A function $f(x_1, x_2, \cdots, x_n)$ in n variables is analytic in an open continuum S in the space of n dimensions if corresponding to each point $(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)$ in S there is a number $\delta > 0$ such that throughout the domain $\mid x_1 - \bar{x}_1 \mid, \mid x_2 - \bar{x}_2 \mid, \cdots, \mid x_n - \bar{x}_n \mid < \delta$, $f(x_1, x_2, \cdots, x_n)$ is representable by a convergent power series in $x_1 - \bar{x}_1, x_2 - \bar{x}_2, \cdots, x_n - \bar{x}_n$. f is analytic in a closed continuum C if it is analytic in an open continuum S containing C.

Here and elsewhere in this paper by "convergence" we understand "absolute convergence." A detailed discussion of the properties of analytic functions of several real variables may be found in Riquier's *Les Systèmes d'Équations aux Dérivées Partielles*, Paris, 1910, Ch. 2, 3.

are supposing that ∇g does not vanish there, each level surface is a regular analytic surface, having a normal at each of its points.

Our interest in the trajectories orthogonal to these surfaces leads us to a consideration of the system of differential equations

$$
\begin{aligned}
dx/dt &= F_1(x, y, z) = g_x/(\nabla g)^2, \\
dy/dt &= F_2(x, y, z) = g_y/(\nabla g)^2, \\
dz/dt &= F_3(x, y, z) = g_z/(\nabla g)^2,
\end{aligned}
\tag{3}
$$

where t is a parameter.

Now since g is analytic in $D - O$ and ∇g does not vanish there the functions F_i are likewise analytic there. We thus have at our disposal the classical existence theorem* on the solutions of (3). This may be stated as follows:

THEOREM A. *Let $P_0(x_0, y_0, z_0)$ denote any fixed point in $D - O$, t_0 an arbitrary real number, and F_{i,P_0} the element of F_i corresponding to the point P_0. There exists a positive number δ_{P_0} and a set of real functions $\phi_{i,P_0}(t, \xi, \eta, \zeta)$, $i = 1, 2, 3$, of t and three independent real variables ξ, η, ζ such that*

(i) *$\phi_{i,P_0}(t, \xi, \eta, \zeta)$ is representable in a power series in $t - t_0$, $\xi - x_0$, $\eta - y_0$, $\zeta - z_0$ which converges for all $|t - t_0|$, $|\xi - x_0|, \cdots$, $|\zeta - z_0| < \delta_{P_0}$*

(ii) *$\phi_{1,P_0}(t_0, \xi, \eta, \zeta) = \xi$, $\phi_{2,P_0}(t_0, \xi, \eta, \zeta) = \eta$, $\phi_{3,P_0}(t_0, \xi, \eta, \zeta) = \zeta$;*

(iii) *$\partial\phi_{i,P_0}/\partial t = F_{i,P_0}(\phi_{1,P_0}, \phi_{2,P_0}, \phi_{3,P_0})$ $(i = 1, 2, 3)$*

for all $|t - t_0|$, $|\xi - x_0|, \cdots$, $|\zeta - 2z_0| < \delta_{P_0}$.
Moreover, the functions $\phi_{i,P_0}(t, x_0, y_0, z_0)$ of t form the only system of integrals of (3) which reduce to x_0, y_0, z_0 respectively when $t = t_0$.

This theorem is fundamental in our analysis.

A set of values of the variables ξ, η, ζ we may regard both as the coördinates of a point (ξ, η, ζ) in the (x, y, z)-space and likewise as defining together with t a point $p(t, \xi, \eta, \zeta)$ an Euclidean 4-space. The first interpretation is of value because the relations

$$
x = \phi_{1,P_0}(t, \xi, \eta, \zeta), \quad y = \phi_{2,P_0}(t, \xi, \eta, \zeta), \quad z = \phi_{3,P_0}(t, \xi, \eta, \zeta),
$$

then may be thought of as representing the trajectories as functions of their

* See for example Goursat, *Cours d'Analyse*, Vol. 2 (1927), p. 18[illegible].

language and the results of the theory of analytic functions of several real variables. We shall have occasion to make use of both interpretations.

Our first theorem is concerned with the latter of these. More particularly it is concerned with the regions of existence in 4-space, and the properties therein, of the analytic functions which the various elements ϕ_{i,P_0} determine. These regions depend, of course, upon the initial point (x_0, y_0, z_0) and the initial value of the parameter t_0. Theorem I is concerned with those solutions for which $t_0 = 1$. It may be seen, however, by an inspection of the theorem that analogous results hold for solutions corresponding to an arbitrary value of t_0.

THEOREM I. *Let R represent the region in 4-space consisting of those points $p(t, \xi, \eta, \zeta)$ such that (ξ, η, ζ) is contained in $D - O$, and t exceeds $1 - g(\xi, \eta, \zeta)$. There exists one and only one set of functions $f_i(t, \xi, \eta, \zeta)$, $i = 1, 2, 3$, having the following properties in R:*

(a) *f_i is analytic in R;*

(b) *$f_1(1, \xi, \eta, \zeta) = \xi$, $f_2(1, \xi, \eta, \zeta) = \eta$, $f_3(1, \xi, \eta, \zeta) = \zeta$, for each point $p(1, \xi, \eta, \zeta)$ in R;*

(c) *the point $P(x, y, z)$, whose coördinates are $x = f_1(t, \xi, \eta, \zeta)$, $y = f_2(t, \xi, \eta, \zeta)$, $z = f_3(t, \xi, \eta, \zeta)$, is in $D - O$ and*

$$
\begin{aligned}
(4) \qquad & g(x, y, z) = t - 1 + g(\xi, \eta, \zeta), \\
& \partial f_i / \partial t = F_i(x, y, z) \qquad (i = 1, 2, 3),
\end{aligned}
$$

for every p in R.

To prove this theorem we first make three preliminary observations relative to the functions ϕ_{i,P_0} referred to in Theorem A. Throughout this proof we shall suppose that the ϕ_{i,P_0} correspond to the initial value of the parameter $t_0 = 1$.

We observe firstly that if δ'_{P_0} be chosen less than* $g(P_0)/4$ and so small that the expansions of the ϕ_{i,P_0} converge for all $|t - 1|$, $|\xi - x_0|, \cdots$, $|\zeta - z_0| < \delta'_{P_0}$, while δ''_{P_0} be chosen so as to satisfy this latter condition as well as to insure that (ξ, η, ζ) is in $D - O$ and $g(\xi, \eta, \zeta)$ exceeds $g(P_0)/2$ whenever $|\xi - x_0|, \cdots, |\zeta - z_0| < \delta''_{P_0}$, then the point $P(x, y, z)$, whose coördinates are

$$
x = \phi_{1,P_0}(t, \xi, \eta, \zeta), \quad y = \phi_{2,P_0}(t, \xi, \eta, \zeta), \quad z = \phi_{3,P_0}(t, \xi, \eta, \zeta),
$$

is in $D - O$ for all $|t - 1| < \delta'_{P_0}$, $|\xi - x_0|, \cdots, |\zeta - z_0| < \delta''_{P_0}$.

---

* Since g is harmonic in $D - O$, positively infinite at O, and zero on d, it is positive at P_0.

For suppose the contrary to be true, that there [illegible] t', ξ', η', ζ' satisfying these conditions and such that the corresponding point P' is not in $D - O$. Then as t varies from one to t', P [illegible] arc. When $t = 1$, P coincides with (ξ', η', ζ'), a point in the δ''_{P_0}-neighborhood of P_0, and thus there is a number τ_1, on the open interval $(1, t')$,* such that when t is on the open interval $(1, \tau_1)$, P is contained in $D - O$, but when $t = \tau_1$, P lies on $d + O$. Hence, because g vanishes continuously on d and becomes positively infinite at O, we have

$$(5) \qquad \lim_{t=\tau_1} g(x, y, z) = 0 \text{ or } \infty,$$

where the appropriate right-hand or left- hand limit is understood.

These relations enable us to reach a contradiction. In fact, we can choose a number τ_2 on the open interval $(1, \tau_1)$ such that for t on the open interval $(1, \tau_2)$, P lies in the δ''_{P_0}-neighborhood of P_0. Now since in this neighborhood

$$F_{i,P_0} = F_i \qquad (i = 1, 2, 3),$$

we must have

$$(\partial\phi_{i,P_0}/\partial t) = F_i(x, y, z) \qquad (i = 1, 2, 3),$$

for t in the open interval $(1, \tau_2)$. Thus, by definition of the F_i,

$$\frac{\partial g}{\partial t} = \frac{\partial g}{\partial x}\,\frac{\partial \phi_{1,P_0}}{\partial t} + \frac{\partial g}{\partial y}\,\frac{\partial \phi_{2,P_0}}{\partial t} + \frac{\partial g}{\partial z}\,\frac{\partial \phi_{3,P_0}}{\partial t} = 1,$$

or

$$g(x, y, z) = t - 1 + g(\xi, \eta, \zeta),$$

for these t. But g being analytic in $D - O$, this equation, because it holds for a small interval in t, must hold as long as P remains within $D - O$, thus for t on the open interval $(1, \tau_1)$. Accordingly,

$$\lim_{t = \tau_1} g(x, y, z) = g(\xi, \eta, \zeta) + \tau_1 - 1,$$

where the limit is taken as above. This quantity is bounded above and exceeds $g(P_0) - 1$, which is greater than zero. This of course contradicts (5).

they hold in a small interval in t for each fixed set of values $|\xi - x_0|, \cdots,$ $|\zeta - z_0| < \delta''_{P_0}$, and the point $(\phi_{1,P_0}, \phi_{2,P_0}, \phi_{3,P_0})$ remains within $D - O$ for all $|t-1| < \delta'_{P_0}$.

Finally, we note that corresponding to every closed subset U of $D - O$, there is a positive number δ_U such that the expansions of the functions ϕ_{i,P_0} in powers of $t-1$, $\xi - x_0, \cdots, \zeta - z_0$ converge for all $|t-1|, \cdots,$ $|\zeta - z_0| < \delta_U$, independently of the position of P_0 in U. In fact, if this were not the case there would exist a sequence of points $\{P_n(x_n, y_n, z_n)\}$, all contained in U, and having a single limit point $\bar{P}(\bar{x}, \bar{y}, \bar{z})$ in U, such that for each n the expansion of one of the functions of the set $\{\phi_{i,P_n}(t, \zeta, \eta, \xi)\}$, associated with P_n, would fail to converge for some $(t_n, \xi_n, \eta_n, \zeta_n)$ such that

$$|t_n - 1| = |\xi_n - x_n| = |\eta_n - y_n| = |z_n - \zeta_n| = 1/n.$$

Accordingly, we could choose a number n' so large that

$$|x_{n'} - \bar{x}|, \cdots, |z_{n'} - \bar{z}| < \delta''_{\bar{P}}/2 \text{ and } 1/n' < \delta''_{\bar{P}}/4,$$

where $\delta''_{\bar{P}}$ has the same meaning with regard to $\bar{P}$ as δ''_{P_0} has with regard to P_0. Now for every system of values $|\xi - x_{n'}|, \cdots, |\zeta - z_{n'}| < \delta''_{Pn'}$, each of the sets of functions $\{\phi_{i,\bar{P}}\}$, $\{\phi_{i,P_n'}\}$ of t, where the $\phi_{i,\bar{P}}$ are the functions associated with $\bar{P}$ by Theorem A, form a system of integrals of (3) and their members reduce to ξ, η, ζ respectively when $t = 1$. Hence, because of the uniqueness of such a set,

$$\phi_{i,\bar{P}} = \phi_{i,P_n'} \qquad (i = 1, 2, 3), \tag{8}$$

for all $|t-1|, \cdots, |\zeta - z'_n| < \delta''_{Pn'}$. From (8) and the fact that expansions of the $\phi_{i,\bar{P}}$ in powers of $t-1$, $\xi - x_{n'}, \cdots, \zeta - z_{n'}$, converge for all $|t-1|, \cdots, |\zeta - z_{n'}| < \delta''_{\bar{P}}/2$, it follows that the functions $\phi_{i,P_n'}$ have this latter property. We then reach a contradiction, for by hypothesis the expansions of the $\phi_{i,P_n'}$ failed to converge for

$$|t_{n'} - 1| = |\xi_{n'} - x_{n'}| = \cdots = |\zeta_{n'} - z_{n'}| = 1/n' < \delta'_{\bar{P}}/4.$$

With the aid of these preliminary remarks we are able to return to the main theorem and prove that with each point $\bar{p}(\bar{t}, \bar{\xi}, \bar{\eta}, \bar{\zeta})$ in R we can associate a neighborhood $C_{\bar{p}}$, defined by inequalities of the type $|t - t| < a$, $|\xi - \bar{\xi}|, \cdots, |\zeta - \bar{\zeta}| < b$, and a set of functions $f_{i\bar{p}}(t, \xi, \eta, \zeta)$, $\mathrm{i} = 1, 2, 3$, such that:

(d) $f_{i,\bar{p}}$ is representable in a power series in $t - \bar{t}, \cdots, \zeta - \bar{\zeta}$ which converges for all $p(t, \xi, \eta, \zeta)$ in $C_{\bar{p}}$;

(e) $f_{1,\bar{p}}(1,\xi,\eta,\zeta)=\xi,\ f_{2,\bar{p}}(1,\xi,\eta,\zeta)=\eta,\ f_{3,\bar{p}}(1,\xi,\eta,\zeta)=\zeta$, in case $\bar{p}$ is contained in the subset R' of R for which $t=1$;

(f) The point $P(x, y, z)$, whose co-ordinates are

$$x=f_{1,\bar{p}}(t,\xi,\eta,\zeta),\ y=f_{2,\bar{p}}(t,\xi,\eta,\zeta),\ z=f_{3,\bar{p}}(t,\xi,\eta,\zeta),$$

is in $D-O$, and

$$g(x,y,z)=t-1+g(\xi,\eta,\zeta) \tag{9}$$
$$\partial f_{i,\bar{p}}/\partial t=F_i(x,y,z) \qquad (i=1,2,3)$$

for all p in $C_{\bar{p}}$;

(g) On each segment $\xi=\text{const.},\cdots,\zeta=\text{const.}$, contained in $C_{\bar{p}}$, the functions $f_{i,\bar{p}}$ are the continuations of functions $\alpha_i(t)$ which are analytic in the neighborhood of $t=1$, reduce to ξ, η, ζ respectively when $t=1$ and have the property that

$$d\alpha_i/dt=F_i(\alpha_1,\alpha_2,\alpha_3) \qquad (i=1,2,3)$$

for t in this neighborhood.

If the truth of these assertions is once established the proof of the main theorem admits immediate conclusion. For, in the first place, the region R, being an open continuum, is one which permits definition of an analytic function therein. In the second place, it follows from (g) that if p_1, p_2 be any two points of R, C_{p_1}, C_{p_2} the neighborhoods, and f_{i,p_1}, f_{i,p_2} the functions associated with them, then

$$f_{i,p_1}(t,\xi,\eta,\zeta)=f_{i,p_2}(t,\xi,\eta,\zeta) \qquad (i=1,2,3), \tag{10}$$

whenever p is in $C_{p_1}\cdot C_{p_2}$. In fact, for every p in $C_{p_1}\cdot C_{p_2}$ the functions $\alpha_i(t)$, corresponding to the three coördinates ξ, η, ζ of p, are uniquely determined and so also are their analytic extensions through real values of t. In the third place, because of (10) and (d), (e), (f) the functions $f_i(t,\xi,\eta,\zeta)$, defined at each point $\bar{p}$ in R as

$$f_i(\bar{t},\bar{\xi},\bar{\eta},\bar{\zeta})=f_{i\bar{p}}(\bar{t},\bar{\xi},\bar{\eta},\bar{\zeta}) \qquad (i=1,2,3),$$

have all the desired properties. Finally, the functions f_i are unique in as much as any set of functions having the properties (a), (b), (c) in R would, because of the uniqueness part of the theorem on the solutions of a differential system, have to agree with the f_i in the neighborhood of some point $\bar{p}$ of R' and thus, because of a well known property of analytic functions, would have to agree with the f_i throughout R.

To obtain the functions $f_{i,\bar{p}}$ for a point $\bar{p}(1,\bar{\xi},\bar{\eta},\bar{\zeta})$ in R' we may apply

[illegible]

our preliminary results directly. In fact, $(\bar{\xi}, \bar{\eta}, \bar{\zeta})$ represents a point P_0 in R, and the functions

$$f_{i,\bar{p}}(t, \xi, \eta, \zeta) = \phi_{i,P_0}(t, \xi, \eta, \zeta) \qquad (i = 1, 2, 3)$$

have the desired properties (d), (e), (f) in the domain $C_{\bar{p}}$ defined by the inequalities

$$|t - 1| < \delta'_{P_0},\ |\xi - \bar{\xi}|, \cdots, |\zeta - \bar{\zeta}| < \delta''_{P_0}.$$

To obtain the functions $f_{i,\bar{p}}$ and establish their properties when $\bar{t} > 1$ one can proceed in much the same way as when $\bar{t} < 1$. We need consider, therefore, only the latter case.

We observe that when $\bar{p}$ is in R and $\bar{t} < 1$ the inequalities

$$g(\bar{\xi}, \bar{\eta}, \bar{\zeta}) > \bar{t} - 1 + g(\bar{\xi}, \bar{\eta}, \bar{\zeta}) = \epsilon' > 0$$

hold. Thus the set $U_{\bar{t}}$ in D determined by the inequalities

$$g(\bar{\xi}, \bar{\eta}, \bar{\zeta}) \geq g \geq \epsilon'$$

is well defined.

Now plainly $U_{\bar{t}}$ is closed and contains only points of $D - O$. Hence there exists a number $\delta_{U_{\bar{t}}} > 0$ so small that expansions in powers of $t - 1$, $\xi - x_0, \cdots, \zeta - z_0$ of the functions ϕ_{i,P_0}, associated with P_0 by Theorem A, converge for all $|t - 1|, \cdots, |\zeta - z_0| < \delta_{U_{\bar{t}}}$, independently of the location of $P_0(x_0, y_0, z_0)$ in $U_{\bar{t}}$.

Let ϵ denote the smaller of the two numbers $\epsilon'/4$, $\delta_{U_{\bar{t}}}$. Let p_0 denote the point whose coördinates are 1, $\bar{\xi}$, $\bar{\eta}$, $\bar{\zeta}$, and let the segment running from p_0 to $\bar{p}$ be divided into n parts by the points of division

$$p_0(t_0, \bar{\xi}, \bar{\eta}, \bar{\zeta}) = p_0(1, \bar{\xi}, \bar{\eta}, \bar{\zeta}),\ p_1(t_1, \bar{\xi}, \bar{\eta}, \bar{\zeta}), \cdots, p_n(t_n, \bar{\xi}, \bar{\eta}, \bar{\zeta}) = \bar{p}(\bar{t}, \bar{\xi}, \bar{\eta}, \bar{\zeta}),$$

in such a way that $0 < t_j - t_{j+1} < \epsilon$, $j = 0, 1, \cdots, n - 1$.

We shall prove by induction that with each point p_j we can associate a set of functions $\{f_{i,p_j}(t, \xi, \eta, \zeta)\}$ and a positive number ϵ_j so small that in the neighborhood B_{p_j} determined by the inequalities $|t - t_j| < \epsilon$, $|\xi - \bar{\xi}|, \cdots$ $|\zeta - \bar{\zeta}| < \epsilon_j$, the functions f_{i,p_j} have the properties (d), (f), (g). When once this is done the proof of the theorem will be complete, for we may identify $C_{\bar{p}}$ with B_{p_n} and $f_{i,\bar{p}}$ with f_{i,p_n}.

We observe firstly that the statement is true for $j = 0$, with ϵ_0 chosen $\leq \delta_{U_{\bar{t}}}$ and so small that (ξ, η, ζ) is in $D - O$ and $g(\xi, \eta, \zeta) > g(\bar{\xi}, \bar{\eta}, \bar{\zeta})$, whenever $|\xi - \bar{\xi}|, \cdots, |\zeta - \bar{\zeta}| < \epsilon_0$. In fact, p_0 is a point R' and with it, as we have seen, there can be associated a neighborhood C_{p_0} and a set of functions having properties (d), (e), (f) in this domain. That these func-

tions have these properties throughout B_{p_0} is readily established from their very definition.

Suppose then that the statement is valid for the point p_j. Let

$$x_0 = f_{1,p}\ (t_{j+1}, \bar{\xi}, \bar{\eta}, \bar{\zeta}),\ y_0 = f_{2,p_j}(t_{j+1}, \bar{\xi}, \bar{\eta}, \bar{\zeta}),\ z_0 = f_{3,p_j}(t_{j+1}, \bar{\xi}, \bar{\eta}, \bar{\zeta}).$$

Then, because the f_{i,p_j} have property (f), we have

$$g(\bar{\xi}, \bar{\eta}, \bar{\zeta}) > g(x_0, y_0, z_0) = t_{j+1} - 1 + g(\bar{\xi}, \bar{\eta}, \bar{\zeta}) \geqq \epsilon'.$$

Thus $P_0(x_0, y_0, z_0)$ is in $U_{\bar{l}}$.

Consider then the functions $\phi_{i,P_0}(t, \xi, \eta, \zeta)$. From the theory of dominant functions it follows that since their expansion in powers of $t-1$, $\xi - x_0, \cdots,$ $\zeta - z_0$ converge for all $|t-1|$, $|\xi - x_0|, \cdots, |\zeta - z_0| < \delta_{U_{\bar{l}}}$, and since the expansions of the $f_{i,p_j}(t_{j+1}, \xi, \eta, \zeta)$ in powers of $\xi - \bar{\xi}$, $\eta - \bar{\eta}$, $\zeta - \bar{\zeta}$ converge for all $|\xi - \bar{\xi}|, \cdots, |\zeta - \bar{\zeta}| < \epsilon_j$ and reduce to x_0, y_0, z_0 respectively when $\xi - \bar{\xi}$, $\eta - \bar{\eta}$, $\zeta - \bar{\zeta}$, there is a positive number $\rho \leqq \epsilon_j$ such that the expansions in powers of $t - t_{j+1}$, $\xi - \bar{\xi}, \cdots, \zeta - \bar{\zeta}$ of the functions

$$f_{i,p_{j+1}}(t, \xi, \eta, \zeta) = \phi_{i,P_0}[t - t_{j+1} + 1, f_{1,p_j}(t_{j+1}, \xi, \eta, \zeta), f_{2,p_j}, f_{3,p}] \quad (i = 1, 2, 3)$$

converge for all $|t - t_{j+1}| < \epsilon$, $|\xi - \bar{\xi}|, \cdots, |\zeta - \bar{\zeta}| < \rho$.

If then we choose $0 < \epsilon_{j+1} \leqq \rho$ and so small that

$$(11) \qquad |f_{1,p_j}(t_{j+1}, \xi, \eta, \zeta) - x_0|, \cdots, |f_{3,p_j}(t_{j+1}, \xi, \eta, \zeta) - z_0| < \delta''_{P_0},$$

when $|\xi - \bar{\xi}|, \cdots, |\zeta - \bar{\zeta}| < \epsilon_{j+1}$, it is readily verified that the functions $f_{i,p_{j+1}}$ have properties (d), (f), (g) in the neighborhood $B_{p_{j+1}}$. Property (d) is, in fact, a direct consequence of the preceding paragraph. Property (f), with the exception of (9), follows from (11) and the definitions of δ''_{P_0} and ϵ. As for (9) we have

$$\begin{aligned} g(f_{1,p_{j+1}}, \cdots, f_{3,p_{j+1}}) &= t - t_{j+1} + 1 - 1 + g[f_{1,p_j}(t_{j+1}, \xi, \eta, \zeta), \cdots, f_{3,p_j}] \\ &= t - t_{j+1} + t_{j+1} - 1 + g(\xi, \eta, \zeta) \\ &= t - 1 + g(\xi, \eta, \zeta). \end{aligned}$$

Finally, in regard to property (g) we observe that for any fixed set of numbers $|\xi - \bar{\xi}|, \cdots, |\zeta - \bar{\zeta}| < \epsilon_{j+1}$ both of the sets of functions $\{f_{i,p_j}\}$, $\{f_{i,p_{j+1}}\}$ of t are analytic in the neighborhood of $t = t_{j+1}$, reduce to $\{f_{i,p_j}(t_{j+1}, \xi, \eta, \zeta)\}$ (a point in D [illegible]) when $t = t_{j+1}$, and form a system of integrals of (3). Since there is at most one such set it follows that

$$f_{i,p_j} = f_{i,p_{j+1}} \qquad (i = 1, 2, 3),$$

[illegible]

[illegible] complete and Theorem I is proved.

2.2. In this section we shall give to the term 'orthogonal trajectory' a meaning which is precise and suited to the problem at hand.

We understand by an orthogonal trajectory one of those curves consisting of the points (x, y, z) determined by the system of formulas

$$x = \psi_1(\tau, x_0, y_0, z_0), \cdots, z = \psi_3(\tau, x_0, y_0, z_0), \tag{12}$$

when $P_0(x_0, y_0, z_0)$ is chosen on the surface $g = 1$ and then held fast, and τ is permitted to vary over the interval $0 \leq \tau < 1$, the functions ψ_i being defined as

$$\psi_i(\tau, x_0, y_0, z_0) = \begin{cases} f_i[(1-\tau)/\tau, x_0, y_0, z_0] & 0 < \tau < 1, \\ 0 & \tau = 0. \end{cases}$$

Conversely, every such curve is an orthogonal trajectory.

In regard to this definition two remarks should be made. In the first place, it should be noted that the functions ψ_i are well defined for $0 \leq \tau < 1$ and P_0 on $g = 1$, for the functions $f_i(t, x_0, y_0, z_0)$ are defined for $\infty > t > 0$ when P_0 is on $g = 1$. Moreover, since

$$g[f_1(t, x_0, \cdots, z_0), \cdots, f_3(t, \cdots, z_0)] = t,$$

when P_0 is on $g = 1$, and g is finite in $d + D - O$, it follows that

$$\lim_{\tau=0} f_i(t, x_0, \cdots, z_0) = 0,$$

or

$$\lim_{\tau=0} f_i[(1-\tau)/\tau, x_0, \cdots, z_0] = 0.$$

Consequently the functions ψ_i are continuous at $\tau = 0$, and the set $\{(x, y, z)\}$ of points (x, y, z) actually constitute a curve in the ordinary sense.

In the second place, it should be pointed out that this definition neither contradicts in any essential way the ordinary notions of an orthogonal trajectory, nor excludes from the family of trajectories any curve which might ordinarily be said to belong there. For on the one hand, every curve represented by (12) certainly cuts the level surfaces orthogonally, and on the other hand, it is readily verified, with the aid of the arguments in 2.52, that any curve, possessing at each of its points a tangent which coincides with the normal to the level surface through that point, coincides throughout its extent with one of the curves represented by (12).* The annexation of the point O to each trajectory is, of course, a slight extension of the ordinary

* Compare Morse, "Relations between the Critical Points of a Real Function of n Independent Variables," *Transactions of the American Mathematical Society*, Vol. 27 (1925), pp. 356-357. In this paper use is made of a differential system analogous to (3) for the study of a certain set of orthogonal trajectories.

[illegible] this correspondence is simply to assign a [illegible] initial point for each trajectory.

[illegible] terms of the point P_t on the surface $u = 1$. In demonstrating the properties of the trajectories and in considering the mapping this representation is, however, by no means the most convenient one. We propose to prove in this section a theorem which shows that the trajectories, at least a certain group of them, can be represented in a very simple way in terms of the directions issuing from O. The proof of this theorem we base on the properties of the solution of the differential equations

$$d\sigma_i/d\mu = \Phi_i(\sigma_1, \sigma_2, \sigma_3) \qquad (i = 1, 2, 3)$$

obtained from the system

$$(13) \qquad dx/d\mu = -|\nabla g|F_1, \; dy/d\mu = -|\nabla g|F_2, \; dz/d\mu = -|\nabla g|F_3$$

in changing over formally to the spherical coördinates $\sigma_1, \sigma_2, \sigma_3$, the colatitude σ_2 being measured from the positive z-axis and the longitude σ_3 from the positive x-axis, counterclockwise with regard to the positive z-axis. We have

$$(14) \qquad \begin{aligned} x &= \theta_1(\sigma_1, \sigma_2, \sigma_3) = \sigma_1 \sin\sigma_2 \cos\sigma_3, \\ y &= \theta_2(\sigma_1, \sigma_2, \sigma_3) = \sigma_1 \sin\sigma_2 \sin\sigma_3, \\ z &= \theta_3(\sigma_1, \sigma_2, \sigma_3) = \sigma_1 \cos\sigma_2. \end{aligned}$$

The functions Φ_i have the form

$$\begin{aligned} \Phi_1(\sigma_1, \sigma_2, \sigma_3) &= (1 - \sigma_1^2 w_{\sigma_1})/K(\sigma_1, \sigma_2, \sigma_3), \\ \Phi_2(\sigma_1, \sigma_2, \sigma_3) &= -w_{\sigma_2}/K, \\ \Phi_3(\sigma_1, \sigma_2, \sigma_3) &= -w_{\sigma_3} \csc^2\sigma_2/K, \end{aligned}$$

where

$$(15) \qquad w(\sigma_1, \sigma_2, \sigma_3) = v(\theta_1, \theta_2, \theta_3),$$

$$K(\sigma_1, \sigma_2, \sigma_3) = [(1 - \sigma_1^2 w_{\sigma_1})^2 + \sigma_1^2 w_{\sigma_2}^2 + \sigma_1^2 w_{\sigma_3}^2 \csc^2\sigma_2]^{1/2}.$$

The functions w, K, when $\sigma_1, \sigma_2, \sigma_3$ are regarded as the coordinates of a point in the Euclidean $(\sigma_1, \sigma_2, \sigma_3)$-space, are, of course, not necessarily [illegible]

there exists a set of functions $\lambda_i(\mu, \alpha, \beta)$, $i = 1, 2, 3$, of μ and two independent variables α, β and a number k_1 so small that in the region U_{k_1}

$$U_{k_1}: \quad |\mu| < k_1, \ \pi/16 < \alpha < 15\pi/16, \ \pi/16 < \beta < 31\pi/16,$$

in the (μ, α, β)-space, the functions λ_i have the following properties:

(i) λ_i is analytic in U_{k_1};

(ii) $\lambda_1(0, \alpha, \beta) = 0$, $\lambda_2(0, \alpha, \beta) = \alpha$, $\lambda_3(0, \alpha, \beta) = \beta$, for $(0, \alpha, \beta)$ in U_{k_1};

(iii) $(\lambda_1, \lambda_2, \lambda_3)$ represents a point in U_h and

$$\partial\lambda_i/\partial\mu = \Phi_i(\lambda_1, \lambda_2, \lambda_3) \qquad (i = 1, 2, 3)$$

for (μ, α, β) in U_{k_1}.

In employing these facts it is not difficult to prove

THEOREM II. *Let Γ, Γ_1 represent the sets of points $q(\tau, \alpha, \beta)$ in the (τ, α, β)-space defined by the inequalities*

$$\Gamma: \quad 0 \leq \tau < 1, \ \pi/8 < \alpha < 7\pi/8, \ \pi/8 < \beta < 15\pi/8,$$
$$\Gamma_1: \quad 0 < \tau < 1, \ \pi/8 < \alpha < 7\pi/8, \ \pi/8 < \beta < 15\pi/8.$$

There exists a set of functions $\phi_i(\tau, \alpha, \beta)$, $i = 1, 2, 3$, having the following properties in Γ, Γ_1:

(a) *$\phi_i(\tau, \alpha, \beta)$ is analytic in Γ_1, and continuous in Γ with $\phi_i(0, \alpha, \beta) = 0$;*

(b) *the point $P(x, y, z)$, whose coördinates are*

$$x = \phi_1(\tau, \alpha, \beta), \ y = \phi_2(\tau, \alpha, \beta), \ z = \phi_3(\tau, \alpha, \beta),$$

is in $D - O$, and

$$g(x, y, z) = 1/\tau - 1$$
$$\partial\phi_i/\partial\tau = -F_i(x, y, z)/\tau^2 \qquad (i = 1, 2, 3),$$

for q in Γ_1;

(c) *the right-hand derivative $(\partial^+\phi_i/\partial\tau^+)_{\tau=0,\, \alpha=\bar{\alpha},\, \beta=\bar{\beta}}$, where*

$$\pi/8 < \alpha < 7\pi/8, \ \pi/8 < \beta < 15\pi/8,$$

exists and is equal to

$$\sin\bar{\alpha}\cos\bar{\beta}, \ \sin\bar{\alpha}\sin\bar{\beta}, \ \cos\bar{\alpha},$$

depending upon whether $i = 1$, 2 or 3;

(d) *the Jacobian*

(16) $$\partial(\phi_1, \phi_2, \phi_3)/\partial(\tau, \alpha, \beta) = \sin\alpha/(\tau^2\nabla g)^2$$

for every q in Γ_1.

It is to be observed that this theorem shows that if ρ be any ray issuing from O which makes an angle $\pi/8 < \alpha < 7\pi/8$ with the positive z-axis and whose projection on the (x, y)-plane makes an angle $\pi/8 < \beta < 15\pi/8$ with the positive x-axis, β being measured counterclockwise with regard to the positive z-axis, then there is a trajectory $C_{\alpha,\beta}$

$$x = \phi_1(\tau, \alpha, \beta),\ y = \phi_2(\tau, \alpha, \beta),\ z = \phi_3(\tau, \alpha, \beta) \tag{17}$$

which has ρ as its forward tangent at O. In fact, by (b) $C_{\alpha,\beta}$ in an orthogonal trajectory, and by (c) it has ρ as its forward tangent at O.

The proof of the theorem we divide into three parts. In (A) the functions ϕ_i are defined and shown to have properties (a), (b). In (B) and (C) these functions are shown to have properties (c) and (d).

(A) We first pass from the functions $\lambda_i(\mu, \alpha, \beta)$ noted above to functions $\bar{\lambda}_i$ of τ, α, β in the following way. We observe that $w(\lambda_1, \lambda_2, \lambda_3)$ is bounded and that $K(\lambda_1, \lambda_2, \lambda_3) > 0$ for (μ, α, β) in U_{k_1}. Thus, since

$$\partial\lambda_1/\partial\mu = [1 - \lambda_1^2 w_{\sigma_1}(\lambda_1, \lambda_2, \lambda_3)]/K(\lambda_1, \lambda_2, \lambda_3),$$

there is a number $0 < k_2 < k_1$ so small that for each (μ, α, β) in the subset

$$U_{k_2}:\ 0 < \mu \leq k_2,\ \pi/8 \leq \alpha \leq 7\pi/8,\ \pi/8 \leq \beta \leq 15\pi/8$$

of U_{k_1}, the function $\partial\lambda_1/\partial\mu > 0$. Thus, for each pair of values

$$\pi/8 \leq \alpha \leq 7\pi/8,\ \pi/8 \leq \beta \leq 15\pi/8, \tag{18}$$

the function λ_1 of μ increases strictly as μ increases from zero to k_2. Accordingly, the function

$$H(\mu, \alpha, \beta) = 1/\lambda_1 + w(\lambda_1, \lambda_2, \lambda_3)$$

is analytic in U_{k_2}. Moreover, since

$$\frac{\partial H}{\partial \mu} = -\frac{1}{\lambda_1^2}\ \frac{\partial \lambda_1}{\partial \mu} + \sum_{i=1}^{3} w_{\sigma_i}\ \frac{\partial \lambda_i}{\partial \mu} = -\frac{K(\lambda_1, \lambda_2, \lambda_3)}{\lambda_1^2}, \tag{19}$$

it follows that, for each fixed pair of values α, β satisfying (18), H increases strictly as μ decreases from k_2 to zero.

Let then $1/\bar{h} - 1$ denote the upper bound of $H(k_2, \alpha, \beta)$ for (α, β) in (18). Then, for every $0 < \tau \leq \bar{h}$ and (α, β) in (18), there is one and only one value $0 < \mu \leq k_2$ for which the equation

$$H(\mu, \alpha, \beta) = 1/\tau - 1$$

is satisfied. This equation thus defines μ as a single valued function

$$\mu = \lambda(\tau, \alpha, \beta)$$

$$U_{\bar{h}}:\quad 0 < \tau < \bar{h},\ \pi/8 < \alpha < 7\pi/8,\ \pi/8 < \beta < 15\pi/8,$$

in the (τ, α, β)-space, is an immediate consequence of the fact that H is analytic and $\partial H/\partial\mu \neq 0$ for (μ, α, β) in U_{k_2}. We have, moreover, (λ, α, β) representing a point in U_{k_2} and

$$H(\lambda, \alpha, \beta) = 1/\lambda_1(\lambda, \alpha, \beta) + w(\lambda_1, \lambda_2, \lambda_3) = 1/\tau - 1 \tag{20}$$

when $q(\tau, \alpha, \beta)$ is in $U_{\bar{h}}$, and

$$\lim_{\tau=0} \lambda(\tau, \alpha, \beta) = 0 \tag{21}$$

for each $\pi/8 < \alpha < 7\pi/8$, $\pi/8 < \beta < 15\pi/8$.

The functions $\bar{\lambda}_i$ are then by definition

$$\bar{\lambda}_i(\tau, \alpha, \beta) = \lambda_i(\lambda, \alpha, \beta) \qquad (i = 1, 2, 3).$$

They have the properties:

(e) $\bar{\lambda}_i$ is analytic in $U_{\bar{h}}$;

(f) $\lim\limits_{\tau=0} \bar{\lambda}_1 = 0$, $\lim\limits_{\tau=0} \bar{\lambda}_2 = \alpha$, $\lim\limits_{\tau=0} \bar{\lambda}_3 = \beta$, for each $\pi/8 < \alpha < 7\pi/8$, $\pi/8 < \beta < 15\pi/8$;

(g) $(\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3)$ is in U_h and

$$\frac{\partial\bar{\lambda}_i}{\partial\tau} = \frac{1}{\tau^2}\,\frac{\partial\lambda_i}{\partial\mu}\,\frac{\bar{\lambda}_1^2}{K(\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3)} = \frac{\bar{\lambda}_1^2\Phi_i(\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3)}{\tau^2 K(\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3)}$$

for each q in $U_{\bar{h}}$.

Property (e) is a consequence of the analyticity of the λ_i, λ. Property (f) follows from (ii) and (21). Property (g) follows from (iii), (19), (20).

We now pass from the $\bar{\lambda}_i$ to the ϕ_i of the theorem. We let

$$\begin{aligned} \bar{\phi}_1(\tau, \alpha, \beta) &= \bar{\lambda}_1 \sin\bar{\lambda}_2 \cos\bar{\lambda}_3, \\ \bar{\phi}_2(\tau, \alpha, \beta) &= \bar{\lambda}_1 \sin\bar{\lambda}_2 \sin\bar{\lambda}_3, \\ \bar{\phi}_3(\tau, \alpha, \beta) &= \bar{\lambda}_1 \cos\bar{\lambda}_2, \end{aligned} \tag{22}$$

and observe that

(h) $\bar{\phi}_i$ is analytic in $U_{\bar{h}}$;

(j) $(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3)$ is a point in D and

$$\begin{aligned} g(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3) &= 1/\bar{\lambda}_1 + v(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3) \\ &= 1/\bar{\lambda}_1 + w(\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3) \\ &= 1/\tau - 1 \end{aligned}$$

when q is in $U_{\bar{h}}$, the first of these equations being a consequence of (22) and the fact that $\bar{\lambda}_1 > 0$ when q is in $U_{\bar{h}}$, the second following from (16), and the last from (20).

We then define

$$\phi_i(\tau, \alpha, \beta) = f_i(t, \xi, \eta, \zeta) \qquad (i = 1, 2, 3), \tag{23}$$

where

$$t = 1/\tau - 2/\bar{h} + 1,$$

$$\xi = \bar{\phi}_1(\bar{h}/2, \alpha, \beta),\ \eta = \bar{\phi}_2(\bar{h}/2, \alpha, \beta),\ \zeta = \bar{\phi}_3(\bar{h}/2, \alpha, \beta),$$

for q in Γ_1, and

$$\phi_i(0, \alpha, \beta) = 0 \qquad (i = 1, 2, 3)$$

for $q(0, \alpha, \beta)$ in Γ.

It is readily verified that the ϕ_i have properties (a), (b). In the first place, the definition is permissible and the functions ϕ_i are analytic in Γ_1 because of (h), Theorem I, and the fact that

$$t = 1/\tau - 2/\bar{h} + 1 > 2 - 2/\bar{h} = 1 - g(\xi, \eta, \zeta).$$

In the second place, by Theorem I, (j) and (23), (ϕ_1, ϕ_2, ϕ_3) represents a point in $D - O$, and

$$\begin{aligned} g(\phi_1, \phi_2, \phi_3) &= g(f_1, f_2, f_3) \\ &= t - 1 + g(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3) \\ &= 1/\tau - 2/\bar{h} + 1 + 2/\bar{h} - 1 \\ &= 1/\tau - 1. \end{aligned}$$

In the third place,

$$(\phi_1^2 + \phi_2^2 + \phi_3^2)^{-1/2} = 1/\tau - 1 - v > 1/\tau - 1$$

so that

$$(\phi_1^2 + \phi_2^2 + \phi_3^2)^{1/2} < \tau/(1 - \tau)$$

and the functions are continuous in Γ. Finally, in the fifth place, by Theorem I and (23)

$$\frac{\partial \phi_i}{\partial \tau} = \frac{\partial f_i}{\partial t}\ \frac{\partial t}{\partial \tau} = -\ \frac{1}{\tau^2}\ F_i(\phi_1, \phi_2, \phi_3) \qquad (i = 1, 2, 3).$$

(B) The ϕ_i thus have properties (a), (b). In order to show that they likewise have property (c) we prove that the $\bar{\phi}_i$ have this property and that

$$\bar{\phi}_i(\tau, \alpha, \beta) = \phi_i(\tau, \alpha, \beta) \qquad (i = 1, 2, 3) \tag{24}$$

when q is in $U_{\bar{h}}$.

Let

$$\bar{\phi}_i(0, \alpha, \beta) = 0.$$

Then in virtue of (j) and (22)

$$\bar{\phi}_1'/\tau = \sin \bar{\lambda}_2 \cos \bar{\lambda}_3 + \bar{\phi}_1[v(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3) + 1].$$

[illegible]

when q is in $U_{\bar{h}}$, it follows by (j) that

$$\lim_{\tau=0} [\bar{\phi}_1(\tau, \alpha, \beta)/\tau] = \sin\alpha\cos\beta.$$

That is,

$$(\partial^+\bar{\phi}_1/\partial\tau^+)_{t=0,\ a=\alpha,\ \beta=\beta} = \sin\alpha\cos\beta.$$

Similarly,

$$(\partial^+\bar{\phi}_2/\partial\tau^+)_{t=0} = \sin\alpha\sin\beta,\quad (\partial^+\bar{\phi}_3/\partial\tau^+)_{t=0} = \cos\alpha.$$

We have then only to verify (24). For this we observe that

$$\begin{aligned}\partial\bar{\phi}_1/\partial\tau = {} & (\partial\bar{\lambda}_1/\partial\tau)\sin\bar{\lambda}_2\cos\bar{\lambda}_3 \\ & + \bar{\lambda}_1\cos\bar{\lambda}_2\cos\bar{\lambda}_3\,(\partial\bar{\lambda}_2/\partial\tau) - \bar{\lambda}_1\sin\bar{\lambda}_2\sin\bar{\lambda}_3\,(\partial\bar{\lambda}_3/\partial\tau),\end{aligned}$$

which by (g) becomes

$$= \bar{\lambda}_1^2(\sin\bar{\lambda}_2\cos\bar{\lambda}_3\Phi_1 + \bar{\lambda}_1\cos\bar{\lambda}_2\cos\bar{\lambda}_3\Phi_2 - \bar{\lambda}_1\sin\bar{\lambda}_2\sin\bar{\lambda}_3\Phi_3)/(\tau^2K^2).$$

Thus, in making use of the formula

$$\begin{aligned}w_{\sigma_1}(\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3) = {} & v_x(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3)\sin\bar{\lambda}_2\cos\bar{\lambda}_3 \\ & + v_y\bar{\lambda}_1\cos\bar{\lambda}_2\cos\bar{\lambda}_3 - v_z\bar{\lambda}_1\sin\bar{\lambda}_2\end{aligned}$$

and the other two of similar nature, we find

$$\partial\bar{\phi}_1/\partial\tau = -F_1(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3)/\tau^2. \tag{25}$$

In the same manner,

$$\partial\bar{\phi}_i/\partial\tau = -F_i(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3)/\tau^2 \qquad (i = 2, 3).$$

Now for $\tau = \bar{h}/2$, we have

$$\phi_i(\bar{h}/2, \alpha, \beta) = \bar{\phi}_i(\bar{h}/2, \alpha, \beta) \qquad (i = 1, 2, 3).$$

This, together with (a), (b), (h), (j) and the uniqueness theorem on the integrals of a different system, is sufficient to establish (24).

(C) To prove that

$$\partial(\phi_1, \phi_2, \phi_3)/\partial(\tau, \alpha, \beta) = \sin\alpha/\tau^2(\nabla g)^2,$$

we first compute $\partial J/\partial\tau$ where,

$$J(\tau, \alpha, \beta) = \tau^2(\nabla g)^2\frac{\partial(\phi_1, \phi_2, \phi_3)}{\partial(\tau, \alpha, \beta)} = -\begin{vmatrix} g_x & \dfrac{\partial\phi_1}{\partial\alpha} & \dfrac{\partial\phi_1}{\partial\beta} \\ g_y & \cdot\cdot & \cdot\cdot \\ g_z & \cdot\cdot & \dfrac{\partial\phi_3}{\partial\beta}\end{vmatrix}.$$

We find upon making use of (b) and the fact that g is harmonic in $D - O$, that

$$\partial J/\partial\tau = 0.$$

Accordingly, J is a function of α, β alone.

Now

$$J = \begin{vmatrix} \phi_1/\tau - \phi_1 - v\phi_1 - v_x r^2 & (\partial\phi_1/\partial\alpha)/r & (\partial\phi_1/\partial\beta)/r \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \phi_3/\tau - \phi_3 - v\phi_3 - v_z r^2 & \cdot & (\partial\phi_3/\partial\beta)/r \end{vmatrix},$$

where $r = (\phi_1^2 + \phi_2^2 + \phi_3^2)^{1/2}$, and we know that for each fixed set of values

(26) $$\pi/8 < \alpha < 7\pi/8,\ \pi/8 < \beta < 15\pi/8,$$

we have

$$\lim_{\tau=0}[\phi_1/\tau] = \sin\alpha\cos\beta,\ \lim_{\tau=0}[\phi_2/\tau] = \sin\alpha\sin\beta,\ \lim_{\tau=0}[\phi_3/\tau] = \cos\alpha,$$
$$\lim_{\tau=0}\phi_i = \lim_{\tau=0}[v\phi_i] = \lim_{\tau=0}[r^2(\nabla v)^2] = 0.$$

Consequently, if we can show that for each fixed pair (α, β) in (26)

(27) $$\lim_{\tau=0}[(\partial\phi_i/\partial\alpha)/r] = \begin{cases} \cos\alpha\cos\beta \\ \cos\alpha\sin\beta, \\ -\sin\alpha \end{cases} \quad \lim_{\tau=0}[(\partial\phi_i/\partial\beta)/r] = \begin{cases} -\sin\alpha\sin\beta & i=1, \\ \sin\alpha\cos\beta & i=2, \\ 0 & i=3, \end{cases}$$

the theorem will be proved, for if these numbers are substituted in their proper places in the above determinant the resulting value is $\sin\alpha$.

All of the limits (27) can be verified by a process similar to that we use for the first.

We have by (22), for an arbitrary pair of values α, β in (26),

$$H(\lambda, \alpha, \beta) = 1/\lambda_1(\lambda, \alpha, \beta) + w(\bar\lambda_1, \bar\lambda_2, \bar\lambda_3) = 1/\tau - 1.$$

Consequently, for all $0 < \tau < \bar h$,

(28) $$\partial\lambda/\partial\alpha[(\partial\lambda_1/\partial\mu) - \lambda_1^2 \sum_{i=1}^{3} w_{\sigma_i}(\bar\lambda_1, \bar\lambda_2, \bar\lambda_3)(\partial\lambda_i/\partial\mu)]$$
$$= -(\partial\lambda_1/\partial\alpha) + \bar\lambda_1^2 \sum_{i=1}^{3} w_{\sigma_i}(\partial\lambda_i/\partial\alpha).$$

Now $\partial\lambda_i/\partial\mu$, $\partial\lambda_i/\partial\alpha$ are bounded for all $0 < \tau < \bar h$, and

(29) $$\lim_{\tau=0}\lambda_i = 0,\ \lim_{\tau=0}(\partial\lambda_1/\partial\alpha) = 0,\ \lim_{\tau=0}(\partial\lambda_1/\partial\mu) = 1,$$

the latter two of these limits resulting from (i), (ii) and the definition of the Φ_i. Thus it follows from (28) that

(30) $$\lim_{\tau=0}(\partial\lambda/\partial\alpha) = 0.$$

The relation (28) may, on the other hand, be written as

$$1 \left(\frac{\partial\lambda_1}{} \frac{\partial\lambda}{} + \frac{\partial\lambda_1}{}\right) = \bar\lambda \sum^{3} w \left(\frac{\partial\lambda_i}{} \frac{\partial\lambda}{} + \frac{\partial\lambda_i}{}\right)$$

[illegible] (29) and (30)

$$\lim_{\tau=0}\left[\left(\frac{\partial\dot{\lambda}_1}{\partial\mu}\ \frac{\partial\lambda}{\partial\alpha}+\frac{\partial\bar{\lambda}_1}{\partial\alpha}\right)\Big/\bar{\lambda}_1\right]=0. \tag{31}$$

This relation enables us to establish the first of the limits in (27). In fact, by (22)

$$\begin{aligned}\frac{1}{r}\ \frac{\partial\phi_1}{\partial\alpha}=\frac{1}{r}\ \frac{\partial\bar{\phi}_1}{\partial\alpha}=\frac{1}{\lambda_1}\left(\frac{\partial\lambda_1}{\partial\mu}\ \frac{\partial\lambda}{\partial\alpha}+\frac{\partial\lambda_1}{\partial\alpha}\right)&\sin\bar{\lambda}_2\cos\bar{\lambda}_3\\ +\left(\frac{\partial\lambda_2}{\partial\mu}\ \frac{\partial\lambda}{\partial\alpha}+\frac{\partial\lambda_2}{\partial\alpha}\right)&\cos\bar{\lambda}_2\cos\bar{\lambda}_3\\ -\left(\frac{\partial\lambda_3}{\partial\mu}\ \frac{\partial\lambda}{\partial\alpha}+\frac{\partial\lambda_3}{\partial\alpha}\right)&\sin\bar{\lambda}_2\sin\bar{\lambda}_3.\end{aligned} \tag{32}$$

But by (ii)

$$\lim_{\tau=0}\ (\partial\lambda_2/\partial\alpha)=1,\quad \lim_{\tau=0}\ (\partial\lambda_3/\partial\alpha)=0.$$

Thus by (31), (32)

$$\lim_{\tau=0}\left(\frac{1}{r}\ \frac{\partial\phi_1}{\partial\alpha}\right)=\cos\alpha\cos\beta,$$

and the theorem is proved.

2.4. The theorem of the preceding section gives us, as we have seen, a representation of a certain group of trajectories, the group being restricted by the bounds upon the variables α, β and thus upon the directions issuing from O. This group may be enlarged and all the directions issuing from O accounted for by a theorem of the same nature as Theorem II. This new theorem will show that if ρ be any ray issuing from O which makes an angle $\pi/8<\alpha'<7\pi/8$, with the positive y-axis and whose projection on the (z,x)-plane makes an angle β' with the negative x-axis, β' being measured counterclockwise with regard to the positive y-axis, then there is a trajectory

$$x=\phi_1'(\tau,\alpha,\beta),\quad y=\phi_2'(\tau,\alpha',\beta'),\quad z=\phi_3'(\tau,\alpha',\beta')$$

having ρ as its forward tangent at O.

The proof of this theorem is based upon the properties of the solutions of the system of differential equations

$$d\sigma_i/d\mu=\Phi_i'(\sigma_1,\sigma_2,\sigma_3)\qquad (i=1,2,3)$$

obtained from the system (13) in changing over to the spherical coördinates $\sigma_1,\sigma_2,\sigma_3$, the colatitude σ_2 being measured in this case from the positive y-axis, and the longitude σ_3 from the negative x-axis, counterclockwise with regard to the positive y-axis. Since, however, the details differ in no essential way from those of Theorem II we shall omit the proof.

THEOREM II′. *Let Γ', Γ_1' represent the sets of points $q'(\tau,\alpha',\beta')$ in the (τ,α',β') space defined by the inequalities*

$$\Gamma': \quad 0 \le \tau \le 1, \quad \pi/8 \le \alpha' \le 7\pi/8, \quad \pi/8 \le \beta' \le 15\pi/8.$$

$$\Gamma_1': \quad 0 < \tau < 1, \quad \pi/8 < \alpha' < 7\pi/8, \quad \pi/8 < \beta' < 15\pi/8.$$

[illegible]

[illegible] properties in Γ', Γ_1'.

(a) *$\phi_i'(\tau, \alpha', \beta')$ is analytic in Γ_1' and continuous in Γ' with* $\phi_i(0, \alpha', \beta') = 0$;

(b) *the point $P(x, y, z)$, whose co-ordinates are*

$$x = \phi_1'(\tau, \alpha', \beta'), \quad y = \phi_2'(\tau, \alpha', \beta'), \quad z = \phi_3'(\tau, \alpha', \beta'),$$

is in $D - O$, and

$$g(x, y, z) = 1/\tau - 1,$$

$$\partial\phi_i'/\partial\tau = -F_i(x, y, z)/\tau^2 \qquad (i = 1, 2, 3)$$

for q' in Γ_1';

(c) *the right-hand derivative* $(\partial^+\phi_i'/\partial\tau^+)_{t=0,\, \alpha'=\bar{\alpha}',\, \beta'=\bar{\beta}'}$, *where*

(33) $$\pi/8 < \bar{\alpha}' < 7\pi/8, \quad \pi/8 < \bar{\beta}' < 15\pi/8$$

exists and is equal to

$$-\sin\bar{\alpha}'\cos\bar{\beta}', \quad \cos\bar{\alpha}', \quad \sin\bar{\alpha}'\sin\bar{\beta}',$$

depending upon whether $i = 1, 2,$ or 3;

(d) *the Jacobian*

$$[\partial(\phi_1', \phi_2', \phi_3')/\partial(\tau, \alpha, \beta)] = \sin\alpha'/[\tau^2(\nabla g)^2]$$

for every q' in Γ_1'.

It is evident from (b) and (c) that the various curves $C'_{\alpha',\beta'}$

$$x = \phi_1'(\tau, \alpha', \beta'), \quad y = \phi_2'(\tau, \alpha', \beta'), \quad z = \phi_3'(\tau, \alpha', \beta')$$

are orthogonal trajectories and have the correct forward tangents at O.

Whether the two families of curves $\{C_{\alpha,\beta}\}$, $\{C'_{\alpha'',\beta''}\}$ include all the trajectories we cannot as yet say. For the [illegible] rays issuing from O which are the forward tangents of [illegible]

To prove the first part of this statement we observe that the points $P^{(n)}$ on T_{P_0} corresponding to the values $\tau = 1 - 1/n$, where n is an integer exceeding 2, have a limit point $\bar{P}$ in $D + d$. $\bar{P}$ cannot, however, lie in D for g is positive and continuous in $D - O$, and becomes infinite continuously at O, while

$$\lim_{n=\infty} g(P_n) = \lim_{n=\infty} [1/(n-1)] = 0.$$

The truth of the second part follows from the facts that T_{P_0} is an analytic arc and contains O, that g is continuous in $D - O$, and finally that

$$g(\psi_1, \psi_2, \psi_3) = (1-\tau)/\tau$$

2.52. *Through every point of $D - O$ there passes one and only one orthogonal trajectory.*

To prove this let us denote by $I_{P'}$ the curve determined by the formulas

$$x = f_1(t, x', y', z'),\ y = f_2(t, \cdots, z'),\ z = f_3(t, \cdots, z')$$

when $P'(x', y', z')$ is fixed in $D - O$, and t is permitted to vary over the interval $\infty > t > 1 - g(x', y', z')$.

Then through each point of $D - O$ there passes an $I_{P'}$, since (x, y, z) coincides with (x', y', z') when $t = 1$. Moreover, in as much as

$$g(x, y, z) = t - 1 + g(x', y', z'),$$

the curve $I_{P'}$ intersects the surface $g = \alpha > 0$, once and only once, namely, when

$$t = \alpha + 1 - g(x', y', z').$$

Further, if $P'(x', y', z')$, $P''(x'', y'', z'')$ be any two points of $D - O$, and $I_{P'}$, $I_{P'}$ their corresponding curves, then $I_{P'}$, $I_{P'}$ coincide throughout if they have a single point $(\bar{x}, \bar{y}, \bar{z})$ in common. In fact, the curve $I_{P'}$ may be defined by the formulas

$$x = f_1[t + 1 - g(x', y', z'), x', y', z'], \cdots,\ z = f_3(t + , \cdots, z'),$$
$$\infty > t > 0,$$

and the curve $I_{P''}$ admits similar representation with x', y', z' replaced by x'', y'', z'', respectively. We have in both cases

$$x = \bar{x},\quad y = \bar{y},\quad z = \bar{z},$$

when $t = g(\bar{x}, \bar{y}, \bar{z}) = \bar{t}$. Thus, as a consequence of the fact that there is at most one system of integrals of (3), analytic in t in the neighborhood of $t = t$, which reduce to $\bar{x}$, $\bar{y}$, $\bar{z}$ when $t = \bar{t}$, it follows that

$$f_i[t + 1 - g(x', y', z'),\ x', y', z'] = f_i\ [t + 1 - g(x'', y'', z''),\ x'', y'', z'']$$
$$(i = 1, 2, 3)$$

in the neighborhood of [illegible] points throughout $\infty > \tau > 0$. Accordingly [illegible] is identical with $I_{P''}$.

From these facts it follows that if P' be any point in $D - O$, there is one and only one trajectory through P'. In fact, the trajectory T_{P_0}, where P_0 is the point of intersection of $I_{P'}$ with $g = 1$, has the desired property, since it contains all the points of I_{P_0} and consequently of $I_{P'}$. On the other hand, no trajectory other than T_{P_0} can contain P'. For if l be any trajectory distinct from T_{P_0} it contains a point P_1, on $g = 1$ distinct from P_0. Now l, except for the point O, is identical with I_{P_1}, but I_{P_1} does not contain P_0 and therefore cannot contain P'. Consequently l cannot contain P'.

2. 53. *Each trajectory is rectifiable over the interval* $0 \leq \tau \leq \frac{1}{2}$.

To prove this we show that the functions $\psi_i(\tau, x_0, y_0, z_0)$ are of bounded variation in this interval for each P_0 on $g = 1$. We have, for any partition of the interval $0 = \tau_1 < \tau_2 \cdots < \tau_n = \frac{1}{2}$

$$\sum_{j=1}^{n-1} | \psi_i(\tau_{j+1}, x_0, y_0, z_0) - \psi_i(\tau_j, x_0, y_0, z_0) | \leq | \psi_i(\tau_1, x_0, y_0, z_0) | + \int_{\tau_1}^{\frac{1}{2}} | \partial\psi_i/\partial\tau | \, d\tau$$

$$\leq | \psi_i(\tau_1, x_0, y_0, z_0) | + \int_{\tau_1}^{\frac{1}{2}} [1/(\tau^2 \, | \nabla g |)] dt,$$

by the definition of ψ_i.

Now

$$(\nabla g)^2 = [1 + 2r^2(\nabla r) \cdot (\nabla v) + r^4(\nabla v)^2]/r^4,$$

and

$$g(\psi_1, \psi_2, \psi_3) = 1/r + v = 1/\tau - 1,$$

so that

$$\tau^4(\nabla g)^2 = [1 + 2r^2(\nabla r) \cdot (\nabla v) + r^4(\nabla v)^2]/(1 + r + rv)^4.$$

This fraction is evidently continuous in the closed subset of D consisting of the origin plus those points where $g \geq 1$, and takes on the value one at the origin. Thus, since $(\nabla g)^2 > 0$ in $D - O$, it follows that

$$\tau^4(\nabla g)^2 > A > 0.$$

2.54. *Each trajectory has a forward tangent at the origin.*

To prove this we need to show that for each P_0 on $g=1$ the right hand derivatives $(\partial^+\psi_i/\partial\tau^+)_{\tau=0}$, or more simply $(\partial\psi_i/\partial\tau)_{\tau=0}$, exist and are not all zero. The forward tangent is then the directed half-line running from O through the point $[(\partial\psi_1/\partial\tau)_{\tau=0}, (\partial\psi_2/\partial\tau)_{\tau=0}, \partial\psi_3/\partial\tau)_{\tau=0}]$.

The proof of the existence is much the same in the three cases $i=1,2,3$. We need therefore consider only the first.

We have in dropping the x_0, y_0, z_0, for the moment, and representing by τ_1, τ_2 two numbers $\frac{1}{2}>\tau_2>\tau_1>0$,

$$[\psi_1(\tau_2)/\tau_2]-[\psi_1(\tau_1)/\tau_1]=[\psi_1(\tau_2)/r_2]-[\psi_1(\tau_1)/r_1] \\ +(1+v_2)\psi_1(\tau_2)-(1+v_1)\psi_1(\tau_1),$$

where r_j and v_j are the values of r and v corresponding to $\tau=\tau_j$. Thus

$$|\,[\psi_1(\tau_2)/\tau_2]-[\psi_1(\tau_1)/\tau_1]\,|\leq\left|\int_{\tau_1}^{\tau_2}[d(\psi_1/r)/d\tau]d\tau\right|+B_1\tau_2,$$

where B_1 is a number independent of τ_1, τ_2 and P_0. Accordingly,

$$\left|\frac{\psi_1(\tau_2)}{\tau_2}-\frac{\psi_1(\tau_1)}{\tau_1}\right|\leq\int_{\tau_1}^{\tau_2}\left|\frac{v_x-x/r(\nabla r)\cdot(\nabla v)}{(\nabla g)^2 r\tau^2}\right|d\tau+B_1\tau_2.$$

Now the quantity

$$|\,v_x-x/r(\nabla r)\cdot(\nabla v)\,|<C,$$

where C is independent of x, y, z in that subset of $D-O$ for which $g\geq 1$. Moreover,

$$(\nabla g)^2 r\tau^2>Ar/\tau^2=A/\tau(1-\tau-\tau v)>A'/\tau,$$

where A' is a positive number independent of $\tau\leq\frac{1}{2}$ and P_0 on $g=1$. Accordingly, there is a number B' independent of P_0 on $g=1$, $0<\tau_1<\tau_2<\frac{1}{2}$, such that

$$|\,\psi_1(\tau_2)/\tau_2-\psi_1(\tau_1)/\tau_1\,|<B'\tau_2.$$

It follows that

$$\lim_{\tau=0}[\psi_1(\tau,x_0,y_0,z_0)/\tau]$$

exists, and exists, be it noted, uniformly with regard to P_0 on $g=1$. Thus $(\partial\psi_1/\partial\tau)_{t=0}$ exists.

That the three derivatives are not all zero at $\tau=0$ follows immediately from the relation

$$\sum_{i=1}^{3}[\psi_i(\tau,x_0,\cdots,z_0)/\tau]^2=(1+rv+r)^2,$$

for the right-hand member of this equality approaches one as τ approaches zero.

2. 55. *Every ray ρ issuing from O is the forward tangent of one and only one orthogonal trajectory.*

That there is one trajectory having the required property we pointed out in Section 2. 4. To show there is at most one it is sufficient to demonstrate that, in using the notation of Sections 2. 3, 2. 4, if $C_{\bar{\alpha},\bar{\beta}}$ or $C_{\bar{\alpha}',\bar{\beta}'}$ has this property then no other trajectory has. Since the reasoning is the same in the two cases we need consider only the former.

We suppose that the trajectory T_{P_0} consisting of points $P_\tau'(x_\tau', y_\tau', z_\tau')$,

$$x_\tau' = \psi_1(\tau, x_0, y_0, z_0), \cdots, z_\tau' = \psi_3(\tau, x_0, y_0, z_0), \qquad 1 \geqq \tau > 0,$$

is distinct from $C_{\bar{\alpha},\bar{\beta}}$ but has ρ as its tangent at O.

We choose α_1, β_1, α_2, β_2 such that

$$\pi/8 < \alpha_1 < \bar{\alpha} < \alpha_2 < 7\pi/8, \; \pi/8 < \beta_1 < \bar{\beta} < \beta_2 < 15\pi/8.$$

We denote by γ_{τ_1} the set of points in Γ_1

$$\gamma_{\tau_1}\colon \quad \tau = \tau_1, \quad \alpha_1 \leqq \alpha \leqq \alpha_2, \quad \beta_1 \leqq \beta \leqq \beta_2,$$

and by D_{τ_1} the points of D corresponding to the points q of γ_{τ_1}. The points of D_{τ_1} are then points of the surface $g = 1/\tau_1 - 1$.

Now for sufficiently small values of τ, say $\tau \leqq \bar{\tau}$, the surface $g = 1/\tau - 1$ is pierced by each ray issuing from O at one and only one point. Hence, in as much as the surface S, $g = 1/\bar{\tau} - 1$, consists of a closed set of points, it follows that S is homeomorphic with a sphere. Thus, since every level surface can be mapped in a 1 : 1 and continuous manner on S by means of the functions $f_i(t, x_0, y_0, z_0)$, every level surface is the homeomorph of a sphere.

Consider then the points E_{τ_1} of D_{τ_1} corresponding to the boundary points e_{τ_1} of γ_{τ_1}. The points E_{τ_1} form a simple closed curve consisting of four analytic arcs, and since the surface $g = 1/\tau_1 - 1$ is homeomorphic with a sphere, this curve divides it into two continua A_{τ_1}, B_{τ_1}, each having E_{τ_1} as its boundary. The set $D_{\tau_1} - E_{\tau_1}$ is certainly contained in one of these continua, say A_{τ_1}, for any two points in $D_{\tau_1} - E_{\tau_1}$ can be joined by a continuous curve not passing through E_{τ_1}. Moreover, every point of A_{τ_1} is a point of $D_\tau - E_\tau$. For if the contrary were true, there would be, because D_τ is [illegible]

$$x' = \phi_1(\tau_1, \alpha_1, \beta_1), \cdots, z' = \phi_3(\tau_1, \alpha_1, \beta_1),$$

contained in $D_\tau - E_\tau$, in every neighborhood of which would lie points of [illegible]

shown to be impossible in making use of the fact that the $\phi_i(\tau, \alpha, \beta)$ are analytic at $\tau_1, \alpha_1, \beta_1$ and have a non-vanishing Jacobian there.

Now P_{τ_1}' cannot belong to D_{τ_1} for any τ_1, for there is at most one trajectory through each point of D_{τ_1} and thus T_{P_0} would have to be one of the curves of the family $\{C_{\alpha,\beta}\}$. But every curve of this family distinct from $C_{\bar{\alpha},\bar{\beta}}$ has a forward tangent at O different from ρ.

Accordingly, P'_{τ_1} is in B_{τ_1} for every τ_1.

But this likewise is impossible. For we observe that the function

$$L(\tau_1, \alpha, \beta) = \sphericalangle\, P_{\tau_1} O \bar{P}_{\tau_1},$$

where P_{τ_1} is the point on D_{τ_1} corresponding to $q(\tau_1, \alpha, \beta)$ in γ_{τ_1} and $\bar{P}_{\tau_1}$ is the point of $C_{\bar{\alpha},\bar{\beta}}$ corresponding to $q(\tau_1, \bar{\alpha}, \bar{\beta})$, is continuous on the closed set e_{τ_1} and consequently takes on its minimum $\omega(\tau_1)$ there. If P_{τ_1}' were contained in B_{τ_1} we should have

$$\sphericalangle\, P_{\tau_1}' O \bar{P}_{\tau_1} > \omega(\tau_1)$$

when $\tau_1 \leq \bar{\tau}$, for otherwise we could pass from P_{τ_1}' to $\bar{P}_{\tau_1}$ by the curve formed from the intersection of the surface $g = 1/\tau_1 - 1$ with the plane determined by the rays $\overrightarrow{OP'_{\tau_1}}$, $\overrightarrow{OP_{\tau_1}}$. It would then follow that

$$\lim_{\tau_1=0} \omega(\tau_1) = 0, \tag{34}$$

for by hypothesis

$$\lim_{\tau_1=0} \sphericalangle P_{\tau_1}' O P_{\tau_1} = 0.$$

But, on the other hand,

$$\lim_{\tau_1=0} L(\tau_1, \alpha, \beta) = L(\alpha, \beta)$$

exists uniformly with regard to q on γ_{τ_1} (Section 2. 54) and

$$L(\alpha, \beta) > 0$$

for all q on γ_{τ_1}. Consequently (34) cannot hold, and therefore P_{τ_1}' cannot be contained in B_{τ_1}.

Thus the property is demonstrated.

3. 1. *The Mapping.* It is evident from the preceding properties of the trajectories that D can be mapped upon the interior D' of a sphere in such a way that the conditions 1 and 2 of the Introduction are satisfied. This may be done, in fact, by erecting in D' a set of rectangular axes x', y', z', with origin at O', the center of D', and taking as the image of the point P in D the point P' in D' lying on the sphere

$$x'^2 + y'^2 + z'^2 = \{\Delta/[1 + g(P)]\}^2,$$

where Δ is the radius of D', and lying on the radius having the same direction angles, referred to the x', y', z'-axes, as the forward tangent at O of the trajectory through P.

That this method of mapping, with the further definition that the image of O be O', satisfies the condition 3, that is, that it yields at $1:1$ correspondence between the points of D and those of D', is not difficult to prove. In fact, on the one hand, since there is one and only one trajectory through each point of D, every such point has a unique image in D' while, on the other hand, every point P' in $D' - O'$ determines uniquely a number τ,

$$x'^2 + y'^2 + z'^2 = (\Delta\tau)^2,$$

and a radius ρ issuing from O, and thus since there is one and only one trajectory having ρ as its forward tangent at O, P' is the image of a unique point P in D.

It can be proved also that the condition 4 is satisfied, that is, that the functions

$$x' = h_1(x, y, z), \quad y' = h_2(x, y, z), \quad z' = h_3(x, y, z),$$

which give the co-ordinates of the point $P'(x', y', z')$, corresponding to the point $P(x, y, z)$, are analytic in $D - O$ and continuous at O.

To show that they are continuous at O we have only to observe that

$$\lim_{p\to O} (x'^2 + y'^2 + z'^2) = \lim_{p\to O} \{\Delta/[1 + g(P)]\}^2$$
$$= 0.$$

To show they are analytic in $D - O$ we observe firstly that when $\Delta\tau$, α, β are regarded as the spherical coördinates of $P'(x', y', z')$, the colatitude α being measured from the positive z'-axis and the longitude β from the positive x'-axis, counterclockwise with regard to the positive z'-axis, then the co-ordinates x, y, z of the point P are given by the formulas

$$x = \phi_1(\tau, \alpha, \beta), \quad y = \phi_2(\tau, \alpha, \beta), \quad z = \phi_3(\tau, \alpha, \beta),$$

provided that P' lies in the subset D_1' of D' determined by the inequalities

$$0 < \tau < 1, \quad \pi/8 < \alpha < 7\pi/8, \quad \pi/8 < \beta < 15\pi/8.$$

This follows from Section 2.3.

Similarly, if $\Delta\tau$, α', β' are regarded as the spherical coördinates of P', the colatitude α' being measured in this case from the positive y'-axis and the longitude β' from the positive z'-axis, counterclockwise with regard [illegible]

$$x = \phi_1'(\tau, \alpha', \beta'), \qquad y = \phi_2'(\tau, \alpha', \beta'), \qquad z = \phi_3'(\tau, \alpha', \beta'),$$

provided that P' lies in the subset D_2' determined by the inequalities

$$0 < \tau < 1, \quad \pi/8 < \alpha' < 7\pi/8, \quad \pi/8 < \beta' < 15\pi/8.$$

Now since

$$D_1' + D_2' = D' - O$$

and τ, α, β are analytic functions of x', y', z' in D_1', and τ, α', β' are analytic functions of these variables in D_2', it follows that the functions

$$x = \bar{h}_1(x', y', z'), \quad y = \bar{h}_2(x', y', z\ ), \quad z = \bar{h}_3(x', y', z'),$$

which give the coördinates of P in terms of those of P' are analytic in $D' - O'$. Moreover we find in employing (d) of Theorems II, II′ that

$$\partial(\bar{h}_1, \bar{h}_2, \bar{h}_3)\ /\partial(x', y', z') = \Delta^3/(\nabla g)^2.$$

Accordingly, the functions h_i, which are the inverses of the $\bar{h}_i$, are analytic in $D - O$.

We summarize these results in the form of a theorem.

THEOREM III. *The region D can be mapped upon the interior D' of a sphere in such a way that the level surfaces of $g(D \mid O, P)$ are carried into spheres concentric with D' and the trajectories orthogonal to these surfaces are carried into radii of D'. When the mapping is done in the manner described above it is* 1:1 *and the coördinates x', y', z' of the point P' in D', corresponding to the point P in D are given by functions*

$$x' = h_1(x, y, z), \quad y' = h_2(x, y, z), \quad z' = h_3(x, y, z),$$

which are analytic in $D - O$ and continuous at O. These functions have the property that

$$\partial(h_1, h_2, h_3)/\partial(x, y, z) = (\nabla g)^2/\Delta^3.$$

3.2. We conclude this paper in observing that the transformations of Section 3.1 generalize completely conformal transformations in the sense that if a plane region is mapped upon the interior of a circle in an analogous manner and the relation between the level curves and their corresponding circles is properly specified then the transformation is necessarily conformal.

The Existence Theorems in the Problem of the Determination of Affine and Metric Spaces by their Differential Invariants.

By Tracy Yerkes Thomas.

Introduction. In a previous article we were concerned with the more formal aspects of the determination of affine and metric spaces by their tensor differential invariants.* We shall now take up the problem from its existence theoretic side. Let us first observe that if the components of a tensor, e. g. the components $A^i{}_{jka}$ or $g_{\alpha\beta,\gamma\delta}$ of a normal or metric tensor, are given originally as analytic functions of the $n(\geqq 2)$ independent variables x^i in the neighborhood of the point $x^i = q^i$, we can by the simple co-ordinate transformation $x^i = y^i + q^i$ transform them into components which are analytic functions of the variables y^i in the neighborhood of the point $y^i = 0$; in fact the direct substitution $x^i = y^i + q^i$ will transform each component $A^i{}_{jka}$ or $g_{\alpha\beta,\gamma\delta}$ with respect to the x^i variables into the corresponding component $A^i{}_{jka}$ or $g_{\alpha\beta,\gamma\delta}$ with respect to the y^i variables. We can therefore assume without loss of generality that the tensor components $A^i{}_{jka}$ or $g_{\alpha\beta,\gamma\delta}$ are analytic functions of the variables y^i in the neighborhood of the point $y^i = 0$. To illustrate more precisely the nature of the problem to be investigated let us consider a definite set of functions $A^i{}_{jka}(y)$ in $n(\geqq 2)$ independent variables y^i; each function $A^i{}_{jka}(y)$ is assumed to be analytic in the neighborhood of the point $y^i = 0$. We wish to investigate the question of whether there exists an affine connection defined by a set of functions $C^i{}_{jk}$ $(= C^i{}_{kj})$ each of which is analytic in the neighborhood of the point $y^i = 0$ such that the $A^i{}_{jka}(y)$ can be derived from the functions $C^i{}_{jk}(y)$ as the components of a normal tensor. An evident requirement is that the functions $A^i{}_{jka}(y)$ should satisfy the complete set of identities

$$\text{(a)} \qquad A^i{}_{jka} = A^i{}_{kja}; \qquad A^i{}_{jka} + A^i{}_{kaj} + A^i{}_{ajk} = 0.$$

In case $n = 2$ we shall show that these conditions are sufficient for the existence of the components $C^i{}_{jk}(y)$ of the affine connection: in fact if the conditions (a) are satisfied the connection $C^i{}_{jk}(y)$ is determined uniquely

---

* "Determination of affine and metric spaces by their differential invariants," *Mathematische Annalen*, Vol. 101 (1929), pp. 713-728. References to [illegible] [illegible] The [illegible] T will be used to refer to this article itself.

within a system of normal co-ordinates y^i. But if $n \geqq 3$ the restrictions imposed by (a) on the functions $A^i{}_{jk\alpha}(y)$ are *not* sufficient to insure the existence of the affine connection $C^i{}_{jk}(y)$. For $n \geqq 3$ it follows that subject to (a) certain of the functions $A^i{}_{jk\alpha}(y)$ can be selected as entirely arbitrary and that others are partially arbitrary in the sense that they can be considered to reduce to arbitrary analytic functions of a number $m(=2, \cdots, n-1)$ of the independent variables y^i for the initial values $y^i = 0$ of the $n-m$ remaining variables. A definite rule can be given for the separation of the functions $A^i{}_{jk\alpha}(y)$ into entirely arbitrary and partially arbitrary functions; this separation being effected the components $C^i{}_{jk}(y)$ of the affine connection are uniquely determined with respect to a system of normal co-ordinates y^i. Analogous remarks apply to the problem of the determination of the functions $g_{\alpha\beta,\gamma\delta}(y)$ in order that there shall exist a symmetric fundamental tensor $h_{\alpha\beta}(y)$ with respect to which the $g_{\alpha\beta,\gamma\delta}(y)$ will be the components of a second extension or metric tensor. Thus the functions $g_{\alpha\beta,\gamma\delta}(y)$ must satisfy the conditions

$$\text{(b)} \qquad g_{\alpha\beta,\gamma\delta} = g_{\beta\alpha,\gamma\delta} = g_{\alpha\beta,\delta\gamma}; \qquad g_{\alpha\beta,\gamma\delta} + g_{\alpha\gamma,\delta\beta} + g_{\alpha\delta,\beta\gamma} = 0$$

which are sufficient to insure the existence of the tensor $h_{\alpha\beta}$ if $n = 2$ but not otherwise; also for $n \geqq 3$ certain of the functions $g_{\alpha\beta,\gamma\delta}$ are entirely arbitrary while certain others reduce to arbitrary functions of a number of the variables y^i for the initial values $y^i = 0$ of the remaining variables. Owing to the analogy between the affine and metric problems these two theories have been developed side by side.

From the standpoint of the theory of differential equations our problem, in the affine case, is to determine a set of analytic functions $A^i{}_{jk\alpha}(y)$ in the most general possible manner so that the two systems of equations (1.1) and (1.2) will be satisfied. We reduce this problem to the problem of finding a set of integrals A_{lm} (the designation A_{lm} is employed to denote certain components $A^i{}_{jk\alpha}$ which remain independent after the conditions (a) have been imposed) of the system (6.1) in which (1) the independent variables y^i are the co-ordinates of a system of normal co-ordinates, and (2) the ★ terms denote polynomials in the components A_{lm} and $C^i{}_{jk}$ such as result by taking the covariant derivative of the tensor $A^i{}_{jk\alpha}$. If the ★ terms were lacking in equations (6.1) these equations would break up into $n-2$ separate systems of equations each of which is of the type treated by König* and their integration could easily be accomplished. On account of

* J. König, "Ueber die Integration simultaner systeme partieller differentialgleichungen mit mehreren unbekannten Functionen," *Mathematische Annalen*, Vol. 23 (1884), pp. 520-526.

the fact that the derivatives of the components C^i_{jk} in the ☆ terms in (b) [illegible] must be related to the components $A^i_{jk\alpha}$ and their derivatives in the process of determining the power series expansions of the $A^i_{jk\alpha}(y)$ we have an [illegible] situation. The more or less elaborate analysis employed has been necessitated by the fact that a situation of this type does not exist in any system of differential equations so far treated in the literature. A similar characterization of the metric problem is possible. In § 1 the conditions on the power series expansions of the functions $A^i_{jk\alpha}$ and $g_{\alpha\beta,\gamma\delta}$ have been given in order that these functions may be the components of a normal or metric tensor respectively. The equations giving these conditions as well as the convergence proofs in § 2 are of importance in the subsequent sections. The finite existence theorems for the functions $A^i_{jk\alpha}$ and $g_{\alpha\beta,\gamma\delta}$ are established in §§ 3, 5, and 6 on the basis of new forms in § 4 of the expressions $N(n, 1+r)$ which gives the number of independent components $A^i_{jk\alpha\beta_1\ldots\beta_r}$ and $N(n, 2+r)$ which gives the number of independent components $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}$ (see M § 1). We have in § 4 derived the new form of the expressions $N(n, p+r)$ for arbitrary values p which can be used in the extension of the results of the present article to normal and metric tensors of higher order. The convergence proofs involved in the finite existence theorems have not been given in the present article as they will appear in a slightly more general form in a later publication.*

1. *General Existence Conditions.* Let us consider that $A^i_{jk\alpha}(y)$ denotes a set of functions of the variable y^i each of which can be expanded about the point $y^i = 0$ in a convergent power series; we shall also assume that the functions $A^i_{jk\alpha}$ satisfy the equations (a). It is desired to identify the variables y^i with the co-ordinates of a system of normal co-ordinates and to choose the functions $A^i_{jk\alpha}$ in such a way, by imposing further necessary conditions, that they will be the components of a normal tensor $A^i_{jk\alpha}(y)$ in this system of co-ordinates. This means that we wish to choose the functions $A^i_{jk\alpha}(y)$ so that the system of equations

(1.1) $$A^i_{jk\alpha}(y) = (\partial C^i_{jk}/\partial y^\alpha) - C^i_{j\alpha,k} - C^i_{\mu j}C^\mu_{k\alpha} - C^i_{j\mu}C^\mu_{k\alpha}$$

is analytic in the neighborhood of the point $y^i = 0$, and such that the equations

$$C^i_{jk}(y) y^j y^k = 0 \tag{1.2}$$

are satisfied identically in the variables y^i. To this end we observe that the functions $A^i_{jka}(y)$ chosen initially so as to satisfy only equations (a) determine an infinite sequence of sets of constants

$$\text{I} \qquad (A^i_{jka\beta_1})_0; \quad (A^i_{jka\beta_1\beta_2})_0; \cdot \cdot \cdot,$$

as explained in M § 4. The constants I determine the formal series for the functions C^i_{jk}, namely

$$C^i_{jk} = (A^i_{jka})_0 y^a + (1/2\,!)(A^i_{jka\beta_1})_0 y^a y^{\beta_1} + \cdot \cdot \cdot, \tag{1.3}$$

where $(A^i_{jka})_0$ has the value $A^i_{jka}(0)$. The series (1.3) evidently satisfy (1.2) since the quantities $(A)_0$ satisfy the complete sets of identities M (1.13). The requirement that (1.3) shall constitute a formal solution of the differential equations (1.1) leads to other conditions on the functions $A^i_{jka}(y)$ than those given by (a). Thus by repeated differentiation of (1.1) and evaluation at the point $y^i = 0$ we obtain equations

$$\begin{aligned} A^i_{jka,\beta_1\ldots\beta_r}(0) &= (2/3)(A^i_{jka\beta_1\ldots\beta_r})_0 \\ &\quad -(1/3)/(A^i_{kaj\beta_1\ldots\beta_r})_0 - (1/3)/(A^i_{ajk\beta_1\ldots\beta_r})_0 + \bigstar, \end{aligned} \tag{1.4}$$

where the $\bigstar$ represents an expression which is quadratic and homogeneous in the components of normal tensors of lower order than those which have been written down explicitly; also

$$A^i_{jka,\beta_1\ldots\beta_r}(0) = \left(\frac{\partial^r A^i_{jka}(y)}{\partial y^{\beta_1} \cdot \cdot \cdot \partial y^{\beta_r}}\right)_{y^i=0}.$$

In (1.4) the right members as well as the left are symmetric in the indices $\beta_1 \cdot \cdot \cdot \beta_r$ arising from the differentiation; no new relations can therefore result from (1.4) as conditions of integrability. Now eliminate the above constants I which stand in the right members of (1.4) by a substitution of the type M (2.10). This will give a system of equations of the general form

$$\begin{aligned} A^i_{jka,\beta_1\ldots\beta_r}(0) &= L\,[A^i_{jka,\beta_1\ldots\beta_r}(0)] \\ &\quad + B[A^i_{jka}(0); \cdot \cdot \cdot; A^i_{jka,\beta_1\ldots\beta_{r-2}}(0)], \end{aligned} \tag{1.5}$$

where as usual L has been used to denote a simple linear sum and B denotes a sum of homogeneous polynomials of its arguments. Equations (1.5) give conditions on the quantities $A^i_{jka,\beta_1\ldots\beta_r}(0)$. If the functions $A^i_{jka}(y)$ satisfy (a) and are furthermore such that the quantities $A^i_{jka,\beta_1\ldots\beta_r}(0)$ will satisfy

[illegible]

in § 2 it will be shown that the normal solution [illegible] (y) given by (1.3) [illegible]. The equations (1.5) therefore give an actual characterization [illegible] the variables y as normal co-ordinates.

THEOREM A. *The set of functions* $A^i_{\ \ \alpha}(y)$, *each of which is analytic* [illegible] *neighborhood of the point* y [illegible] *the components of a* [illegible] *tensor in a system of normal co-ordinates y^i if, and only if, the functions $A^i_{\ \ \alpha}(y)$ satisfy (a) and are furthermore such that the coefficients $A^i_{\ \ j,\alpha,\beta_1\ldots\beta_r}(0)$ of their power series expansions about the point $y^i = 0$ satisfy the sequence of equations* (1.5).

The precise form of the equations (1.5) for $r = 1$ can be obtained by eliminating the quantities $A^i_{\ \ j\kappa\sigma\beta}$ in the right members of $M(2.1)$ by means of $M(2.8)$. This gives

$$(1.6) \qquad 3A^i_{\ \ jk\alpha\beta} = 3A^i_{\ \ j\alpha\beta,k} + 2A^i_{\ \ j\beta\alpha,k} + 2A^i_{\ \ k\beta\alpha,j} + A^i_{\ \ \alpha\beta j,k} + A^i_{\ \ \alpha\beta k,j}$$

as the conditions to be satisfied by the quantities

$$A^i_{\ \ jk\alpha,\beta} = A^i_{\ \ jk\alpha,\beta}(0).$$

Now consider a set of constants $(g_{\alpha\beta})_0 = (g_{\beta\alpha})_0$, such that the determinant $|(g_{\alpha\beta})_0| \neq 0$, and a set of functions $g_{\alpha\beta,\gamma\delta}(y)$ each of which is analytic in the neighborhood of the point $y^i = 0$. Let us state the conditions that must be satisfied by the functions $g_{\alpha\beta,\gamma\delta}(y)$ so that they will be the components of a metric tensor resulting from a fundamental tensor with symmetric components $h_{\alpha\beta}(y)$ in a system of normal co-ordinates y^i such that $h_{\alpha\beta}(0) = (g_{\alpha\beta})_0$ at the origin of the system. In other words we wish to find conditions on the functions $g_{\alpha\beta,\gamma\delta}(y)$ so that there will exist a solution of the system of differential equations

$$g_{\alpha\beta,\gamma\delta}(y) \ldots [illegible]$$

[illegible]

identically in the variables y^i. These latter equations can be shown to be equivalent to (1.2) in which the $C^i_{\alpha\beta}(y)$ are Christoffel symbols, i. e.

$$C^i_{\alpha\beta}(y) = \tfrac{1}{2}h^{i\sigma}[(\partial h_{\alpha\sigma}/\partial y^\beta) + (\partial h_{\sigma\beta}/\partial y^\alpha) - (\partial h_{\alpha\beta}/\partial y^\sigma)].$$

It is now expedient however to use (1.8) rather than (1.2) to characterize the variables y^i as normal co-ordinates. On the basis of the Christoffel symbols $C^i_{\alpha\beta}$ the quantities $C^i_{\alpha\beta\gamma}$ in (1.7) as well as the quantities $C^i_{\alpha\beta\ldots\epsilon}$ in general which we shall need later are defined by the equations given in M § 1. We first assume that the functions $g_{\alpha\beta,\gamma\delta}(y)$ satisfy the complete set of identities (b) as this is a necessary condition. The process described in M § 4 will then determine the sets of constants

$$\text{II} \qquad (g_{\alpha\beta,\gamma\delta\epsilon_1})_0; \qquad (g_{\alpha\beta,\gamma\delta\epsilon_1\epsilon_2})_0; \cdots$$

in such a way that the formal series for $h_{\alpha\beta}$, namely

$$(1.9) \qquad h_{\alpha\beta} = (g_{\alpha\beta})_0 + (1/2!)(g_{\alpha\beta,\gamma\delta})_0 y^\gamma y^\delta + (1/3!)(g_{\alpha\beta,\gamma\delta\epsilon_1})_0 y^\gamma y^\delta y^{\epsilon_1} + \cdots,$$

will satisfy (1.8); in the series (1.9) the constant $(g_{\alpha\beta,\gamma\delta})_0$ has the value $g_{\alpha\beta,\gamma\delta}(0)$. Now differentiate (1.7) repeatedly, evaluate at the point $y^i = 0$ and eliminate the quantities $(g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r})_0$ in the right members of the resulting equations by a substitution of the type M(2.16). This gives conditions analogous to (1.5), namely

$$(1.10) \qquad g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}(0) = L\,[g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}(0)] + Q[(g_{\alpha\beta})_0;\ g_{\alpha\beta,\gamma\delta}(0);\cdots;\ g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_{r-2}}(0)],$$

which must be satisfied by the quantities

$$g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}(0) = \left(\frac{\partial^r g_{\alpha\beta,\gamma\delta}(y)}{\partial y^{\epsilon_1}\cdots\partial y^{\epsilon_r}}\right)_{y^i=0}$$

in order that the series (1.9) should constitute a formal solution (1.7). As it will be shown in § 2 that this formal solution (1.8) converges we have the following theorem.

THEOREM B. *Given a set of constants $(g_{\alpha\beta})_0 = (g_{\beta\alpha})_0$, such that the determinant $|(g_{\alpha\beta})_0| \neq 0$, and a set of functions $g_{\alpha\beta,\gamma\delta}(y)$ each of which is analytic in the neighborhood of the point $y^i = 0$. There will then exist a fundamental tensor having components $h_{\alpha\beta}(y) = h_{\beta\alpha}(y)$ in a system of normal co-ordinates y^i such that* (1) *$h_{\alpha\beta}(0) = (g_{\alpha\beta})_0$, and* (2) *the functions $g_{\alpha\beta,\gamma\delta}(y)$ are derivable from the $h_{\alpha\beta}(y)$ as the components of a metric tensor if, and only if, the functions $g_{\alpha\beta,\gamma\delta}(y)$ satisfy equations* (b) *and are furthermore*

such that the coefficients $g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}(0)$ *of their power series expansions about the point* $y^i = 0$ *satisfy the sequence of equations* (1.10).

To deduce equations corresponding to (1.6) differentiate (1.7) and evaluate at the point $y^i = 0$. This gives

$$6 g_{\alpha\beta,\gamma\delta,\epsilon} = 4 g_{\alpha\beta,\gamma\delta\epsilon} + g_{\alpha\epsilon,\beta\gamma\delta} + g_{\beta\epsilon,\alpha\gamma\delta} + 2 g_{\gamma\delta,\alpha\beta\epsilon}, \tag{1.11}$$

where $g_{\alpha\beta,\gamma\delta\epsilon}$ denotes the third derivative of $h_{\alpha\beta}$ with respect to y^γ, y^δ, y^ϵ evaluated at the point $y^i = 0$, and

$$g_{\alpha\beta,\gamma\delta,\epsilon} = \left(\frac{\partial g_{\alpha\beta,\gamma\delta}(y)}{\partial y^\epsilon}\right)_{y^i=0}.$$

Eliminating the quantities $g_{\alpha\beta,\gamma\delta\epsilon}$ in the right members of (1.10) by M(2.18) we deduce

$$2 g_{\alpha\beta,\gamma\delta,\epsilon} = g_{\alpha\beta,\delta\epsilon,\gamma} + g_{\alpha\beta,\gamma\epsilon,\delta} + g_{\beta\epsilon,\gamma\delta,\alpha} + g_{\alpha\epsilon,\gamma\delta,\beta}. \tag{1.12}$$

Equations (1.6) and (1.12) can also be considered to give relations between the covariant derivatives $A^i_{jk\alpha,\beta}$ and $g_{\alpha\beta,\gamma\delta,\epsilon}$ respectively which must hold throughout the neighborhood of the origin of the normal co-ordinate system. This more general interpretation of (1.6) and (1.12) will be of importance later.

2. *Convergence proofs.* A proof of the convergence of the series (1.3) and (1.9) can be made by the method involving the use of dominant functions, i. e. the *Calcul des limites* of Cauchy. Let us first prove the convergence of the series (1.3) where it is to be understood that this series is determined as above explained by a set of functions $A^i_{jk\alpha}(y)$ which satisfy the conditions stated in Theorem A. It can be shown readily that the function $\mathfrak{C}^i_{jk}$ defined as a solution of the system of differential equations

$$\partial \mathfrak{C}^i_{jk}/\partial y^\alpha = 4 \mathfrak{A}^i_{jk\alpha} + 2(\mathfrak{C}^i_{\mu k}\mathfrak{C}^\mu_{j\alpha} + \mathfrak{C}^i_{\mu\alpha}\mathfrak{C}^\mu_{jk}) \tag{2.1}$$

such that $\mathfrak{C}^i_{jk} = 0$ when $y^1 = \cdots = y^n = 0$, is a dominant function for the connection C^i_{jk} given by (1.3); in (2.1) the function $\mathfrak{A}^i_{jk\alpha}$ is a dominant function for $A^i_{jk\alpha}(y)$. Equations (2.1) can in fact be reduced to a very simple form by putting

$$\mathfrak{C}^i_{jk} = \Phi; \qquad \mathfrak{A}^i_{jk\alpha} = F,$$

where F, defined by the equation

for suitable positive constants M and ρ, is a dominant function for any of the functions $A^i{}_{jka}$. With the above substitutions (2.1) becomes

$$(2.3) \qquad d\Phi/dy = 4F + 4n\Phi^2,$$

where $y = y^1 + \cdots + y^n$. The existence of a solution Φ of the equation (2.3) which is analytic in the neighborhood of $y = 0$, and such that $\Phi = 0$ for $y = 0$, results from the well known theorem of partial differential equations. To show that the function Φ satisfying these initial conditions dominates the series (1.3) for the affine connection $C^i{}_{jk}$ let us differentiate (2.1) repeatedly and evaluate the resulting equations at the point $y^i = 0$. We thus obtain expressions of the general form

$$(2.4) \qquad \begin{aligned} \Phi^i{}_{jka\beta_1\ldots\beta_r} = {} & 4F^i{}_{jka\beta_1\ldots\beta_r} \\ & + 2(\partial^r/\partial y^{\beta_1}\cdots\partial y^{\beta_r})(\mathfrak{C}^i{}_{\mu k}\mathfrak{C}^\mu{}_{ja} + \mathfrak{C}^i{}_{\mu a}\mathfrak{C}^\mu{}_{jk})_{y^i=0}, \end{aligned}$$

where $\Phi^i{}_{jka\beta_1\ldots\beta_r}$ and $F^i{}_{jka\beta_1\ldots\beta_r}$ denote the derivatives of $\partial\mathfrak{C}^i{}_{jk}/\partial y^a$ and $\mathfrak{A}^i{}_{jka}$ with respect to $y^{\beta_1}, \cdots, y^{\beta_r}$ evaluated at $y^i = 0$. Assuming that

$$(2.5) \qquad \Phi^i{}_{jka\beta_1\ldots\beta_s} > |(A^i{}_{jka\beta_1\ldots\beta_s})_0|$$

for $s < r (\geqq 2)$ we can show that for $s = r$ these inequalities are likewise satisfied. Compare equations (2.4) with equations

$$(2.6) \qquad A^i{}_{jka\beta_1\ldots\beta_r} = A^i{}_{jak\beta_1\ldots\beta_r} + A^i{}_{jka,\beta_1\ldots\beta_r} - A^i{}_{jak,\beta_1\ldots\beta_r} + \bigstar,$$

(see M § 2) in which $A^i{}_{jka}$ in the $\bigstar$ term has the value $A^i{}_{jka}(0)$ and the quantities $A^i{}_{jka\beta_1\ldots\beta_r}$ and $A^i{}_{jka,\beta_1\ldots\beta_r}$ are considered to represent the constants $(A^i{}_{jka\beta_1\ldots\beta_r})_0$ and $A^i{}_{jka,\beta_1\ldots\beta_r}(0)$ for any value of $r(\geqq 1)$. This comparison shows that

$$2(\partial^r/\partial y^{\beta_1}\cdots\partial y^{\beta_r})(\mathfrak{C}^i{}_{\mu k}\mathfrak{C}^\mu{}_{ja} + \mathfrak{C}^i{}_{\mu a}\mathfrak{C}^\mu{}_{jk})_{y^i=0}$$

is greater than twice the absolute value of the $\bigstar$ terms in (2.6) in virtue of (2.5) for $s < r$; also $4F^i{}_{jka\beta_1\ldots\beta_r}$ is greater than twice the absolute value of the difference $A^i{}_{jka,\beta_1\ldots\beta_r} - A^i{}_{jak,\beta_1\ldots\beta_r}$ in (2.6) because of the dominating property of the function F. Hence

$$(2.7) \qquad \Phi^i{}_{jka\beta_1\ldots\beta_r} > 2\,|(A^i{}_{jka\beta_1\ldots\beta_r})_0 - (A^i{}_{jak\beta_1\ldots\beta_r})_0|.$$

Since the $\Phi^i{}_{jka\beta_1\ldots\beta_r}$ are symmetric with respect to all indices we also have

$$(2.8) \qquad \Phi^i{}_{jka\beta_1\ldots\beta_r} > 2\,|(A^i{}_{jak\beta_1\ldots\beta_r})_0 - (A^i{}_{a\beta_1 jk\ldots\beta_r})_0|.$$

Adding (2.7) and (2.8) we obtain

$$\Phi^i{}_{jka\beta_1\ldots\beta_r} > |(A^i{}_{jka\beta_1\ldots\beta_r})_0 - (A^i{}_{a\beta_1 jk\ldots\beta_r})_0|.$$

Hence $\Phi^i{}_{jk\alpha\beta_1\ldots\beta_r}$ is greater than the absolute value of the difference of $(A^i{}_{jk\alpha\beta_1\ldots\beta_r})_0$ and $(A^i{}_{pq\sigma\tau_1\ldots\tau_r})_0$ where $pq\sigma\tau_1\cdots\tau_r$ is any permutation of $jk\alpha\beta_1\cdots\beta_r$; we can therefore write

$$\Phi^i{}_{jk\alpha\beta_1\ldots\beta_r} > \left| (A^i{}_{jk\alpha\beta_1\ldots\beta_r})_0 - (A^i{}_{pq\sigma\tau_1\ldots\tau_r})_0 \right|. \tag{2.9}$$

Now form all permutations of the indices $pq\sigma\tau_1\cdots\tau_r$ and add together the $\mathfrak{N}$ corresponding inequalities (2.9) obtaining

$$\mathfrak{N}\Phi^i{}_{jk\alpha\beta_1\ldots\beta_r} > \left| \mathfrak{N}(A^i{}_{jk\alpha\beta_1\ldots\beta_r})_0 - \sum (A^i{}_{pq\sigma\tau_1\ldots\tau_r})_0 \right|.$$

But the sum in the right member of this inequality vanishes by M(1.13). Hence we obtain the inequality (2.5) for $s = r(\geqq 2)$ which we wished to prove. The fact that (2.5) holds for $s = 1$ is proved directly by observing that for this case the ★ terms in (2.6) and the second set of terms in (2.4) vanish; we are thus led to the inequality (2.7) for $r = 1$ and hence to (2.5) for $s = 1$. The function $\mathfrak{C}^i{}_{jk}$ therefore dominates the series (1.3) for the affine connection $C^i{}_{jk}$ with the result that (1.3) must converge.

To prove the convergence of the series (1.9) we set up the system of differential equations

$$\partial^2\mathfrak{H}_{\alpha\beta}/\partial y^\gamma \partial y^\delta = 4\mathfrak{G}_{\alpha\beta\gamma\delta} + 4\sum \left[(\partial\mathfrak{H}_{\sigma\alpha}/\partial y^\beta)\mathfrak{R}^\sigma{}_{\gamma\delta} + \mathfrak{H}_{\sigma\tau}\mathfrak{R}^\sigma{}_{\alpha\beta}\mathfrak{R}^\tau{}_{\gamma\delta}\right], \tag{2.10}$$

where $\mathfrak{G}_{\alpha\beta\gamma\delta}$ is a dominant function for $g_{\alpha\beta,\gamma\delta}(y)$ and the summation $\sum$ denotes the sum of all terms that can be formed from the term

$$(\partial\mathfrak{H}_{\sigma\alpha}/\partial y^\beta)\mathfrak{R}^\sigma{}_{\gamma\delta} + \mathfrak{H}_{\sigma\tau}\mathfrak{R}^\sigma{}_{\alpha\beta}\mathfrak{R}^\tau{}_{\gamma\delta} \tag{2.11}$$

by taking all permutations of the indices $\alpha\beta\gamma\delta$. Also the expression $\mathfrak{R}^i{}_{\alpha\beta}$ in the term (2.11) is defined by

$$\mathfrak{R}^i{}_{\alpha\beta} = \mathfrak{L}^{i\sigma}\left[(\partial\mathfrak{H}_{\sigma\beta}/\partial y^\alpha) + (\partial\mathfrak{H}_{\sigma\alpha}/\partial y^\beta) + (\partial\mathfrak{H}_{\alpha\beta}/\partial y^\sigma)\right],$$

where $\mathfrak{L}^{i\sigma}$ is a dominant function for $\mathfrak{H}^{i\sigma}$. Now replace $\mathfrak{H}_{\alpha\beta}$ by Ψ and take $\Psi_0 > |(g_{\alpha\beta})_0|$ (absolute value); also take

$$\mathfrak{G}_{\alpha\beta\gamma\delta} = \frac{M}{1 - (y/\rho)}; \qquad \mathfrak{L}^{i\sigma} = \frac{M'}{1 - (\Psi - \Psi_0)/\rho'},$$

where $y = y^1 + \cdots + y^n$ and M, M', ρ, ρ' are suitable positive constants. The equations (2.10) will then reduce to a single differential equation as in the case previously considered. This differential equation possesses a

such that $\Psi(0)=\Psi_0$ and $d\Psi(y)/dy=0$ for $y=0$. The function $\Psi(y)$ so determined dominates the series (1.9). To show this we assume that

$$\Psi_{\alpha\beta\gamma\delta\epsilon_1\ldots\epsilon_s} > |(g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_s})_0| \tag{2.12}$$

for $s < r(\geqq 2)$ where the quantity $\Psi_{\alpha\beta\gamma\delta\epsilon_1\ldots\epsilon_s}$ is equal to the derivative of $\mathfrak{H}_{\alpha\beta}$ with respect to $y^\gamma\cdots y^{\epsilon_s}$ evaluated at $y^i=0$. We shall show that (2.12) is also true for $s=r$ which will prove the dominating property of the function $\Psi(y)$. Due to the fact that $\mathfrak{H}_{\alpha\beta}$ is to be taken equal to the function $\Psi(y)$ it follows that $\Psi_{\alpha\beta\gamma\delta\epsilon_1\ldots\epsilon_s}$ is symmetric with respect to all its indices. To prove (2.12) for $s=r$ we differentiate (2.10) repeatedly r times, then evaluate at $y^i=0$ and compare the resulting equations with the equations

$$\begin{aligned} &g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r} - g_{\alpha\gamma,\beta\delta\epsilon_1\ldots\epsilon_r} + g_{\delta\gamma,\beta\alpha\epsilon_1\ldots\epsilon_r} - g_{\delta\beta,\gamma\alpha\epsilon_1\ldots\epsilon_r} \\ &= g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r} - g_{\alpha\gamma,\beta\delta,\epsilon_1\ldots\epsilon_r} + g_{\delta\gamma,\beta\alpha,\epsilon_1\ldots\epsilon_r} - g_{\delta\beta,\gamma\alpha,\epsilon_1\ldots\epsilon_r} + \bigstar \end{aligned} \tag{2.13}$$

[equations $M(2.14)$], in which the $g_{\alpha\beta,\gamma\delta}$ appearing in the $\bigstar$ terms has the value $g_{\alpha\beta,\gamma\delta}(0)$ and the quantities $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}$ and $g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}$ have the values $(g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r})_0$ and $g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}(0)$ respectively for any value of $r(\geqq 1)$. It is clear that

$$\begin{aligned} \Psi_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r} > |&(g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r})_0 - (g_{\alpha\gamma,\beta\delta\epsilon_1\ldots\epsilon_r})_0 \\ &+ (g_{\gamma\delta,\beta\alpha\epsilon_1\ldots\epsilon_r})_0 - (g_{\delta\beta,\gamma\alpha\epsilon_1\ldots\epsilon_r})_0|. \end{aligned} \tag{2.14}$$

We now introduce the operations P and P' defined in M § 2. It will be recalled that the operation P on a term such as $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}$ effects the summation of all terms obtainable from $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}$ by permutting the indices $\delta\epsilon_1\cdots\epsilon_r$ cyclically while the remaining indices $\alpha\beta\gamma$ are held fixed; similarly the operation P' on the term $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}$ effects the summation of the terms obtainable from $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}$ by taking all permutations of the indices $\gamma\delta\epsilon_1\cdots\epsilon_r$. First operating on both members of (2.14) by P we have

$$(r+1)\Psi_{\alpha\beta\gamma\delta\epsilon_1\ldots\epsilon_r} > (r+2)|(g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r})_0 - (g_{\alpha\gamma,\beta\delta\epsilon_1\ldots\epsilon_r})_0|.$$

We next operate on both members of these latter inequalities by P' which gives

$$(r+1)\Psi_{\alpha\beta\gamma\delta\epsilon_1\ldots\epsilon_r} > (r+3)|(g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r})_0|.$$

This proves (2.12) for $s=r(\geqq 2)$. If $s=1$ we obtain the inequality (2.14) for $r=1$ since the $\bigstar$ terms in (2.10) and (2.13) then vanish; operating on (2.14) for $r=1$ as above we obtain (2.12) for $s=1$. The function $\Psi(y)$ is therefore a dominant function for the series (1.9) and therefore the series (1.9) converges within a sufficiently small neighborhood of the point $y^i=0$.

3. *Special Case of Two Variables.* It is of interest to investigate the question of whether the conditions given by (1. 5) and (1. 10) on the derivatives of the functions $A^i_{jk\alpha}(y)$ and $g_{\alpha\beta,\gamma\delta}(y)$ at $y^i = 0$ are independent of those obtainable from the complete sets of identities of these functions, i. e. conditions of the type M(1. 15) and M(1. 14). The simplest case to consider is that of the functions $A^i_{jk\alpha}(y)$ or $g_{\alpha\beta,\gamma\delta}(y)$ for $n = 2$, i. e. for two variables. First take the functions $A^i_{jk\alpha}(y)$ and suppose that they satisfy the complete set of identities (a). The number $N(2,1)$ defined in M § 1, and which is equal to 4, gives the number of independent components $A^i_{jk\alpha}$. Now if (a) gave all the conditions on the function $A^i_{jk\alpha}(y)$, the number of independent components of the extension $A^i_{jk\alpha,\beta_1\ldots\beta_r}$ would be $4K(2,r)$ or $4(r+1)$. But actually the number of independent quantities $A^i_{jk\alpha,\beta_1\ldots\beta_r}$ is given by $N(2,r+1)$* which is equal to $4(r+1)$. Hence it follows that the equation (1. 5) for $n = 2$ can not produce additional conditions on the quantities $A^i_{jk\alpha,\beta_1\ldots\beta_r}$ over those which result from (a) by extension.

There is of course an analogous result for the functions $g_{\alpha\beta,\gamma\delta}(y)$ in two variables. Here the number of independent components $g_{\alpha\beta,\gamma\delta}(y)$ is $N(2,2)$ or 1. Then, since $K(2,r)$ and $N(2,r+2)$ each have the value $r+1$, it follows that (b) gives all the functional conditions on the components $g_{\alpha\beta,\gamma\delta}(y)$; the conditions (1. 10) for $n = 2$ must therefore be satisfied identically in this case as a consequence of (b) and the conditions obtainable from (b) by extension.

THEOREM C. *For $n = 2$ the functions $A^i_{jk\alpha}(y)$, each of which is analytic in the neighborhood of the point $y^i = 0$, will be the components of a normal tensor in a system of normal co-ordinates if, and only if, the functions*

---

* Since the components $A^i_{jk\alpha,\beta}$ are expressible in terms of the components $A^i_{jk\alpha\beta}$ by M(2. 1) and the components $A^i_{jk\alpha\beta}$ are expressible in terms of the components $A^i_{jk\alpha,\beta}$ by M(2. 8), it follows that the number of independent components $A^i_{jk\alpha,\beta}$ is equal to the number of independent components $A^i_{jk\alpha\beta}$. More generally if we consider the values of the components $A^i_{jk\alpha\beta_1\ldots\beta_s}$ and $A^i_{jk\alpha,\beta_1\ldots\beta_s}$ for $s < r$ to be determined then the number of independent components $A^i_{jk\alpha,\beta_1\ldots\beta_r}$ is equal to the number of independent components $A^i_{jk\alpha\beta_1\ldots\beta_r}$ on account of the equations

$$A^i_{jk\alpha\beta_1\ldots\beta_r} \; [illegible] \; A^i_{jk\alpha\beta_1\ldots\beta_r} - A^i_{[illegible]\ldots\beta_r} + \star$$

(see M § 2) and the equations M(2. 10). Similarly the number of independent components $g_{\alpha\beta\gamma\delta\epsilon_1\ldots\epsilon_r}$ is equal to the number of independent components $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}$ when the values (1) of the components $g_{\alpha\beta}$, and (2) of the components $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_s}$ and $g_{\alpha\beta\gamma\delta\epsilon_1\ldots\epsilon_s}$ for $s < r$ are previously determined. In this sense we can say [illegible]

[illegible] $N(n, 2, r)$ as given by (1. 13) in § 1.

$A^i{}_{jk\alpha}(y)$ *satisfy* (*a*). *Similarly, for* $n=2$ *the functions* $g_{\alpha\beta,\gamma\delta}(y)$, *each of which is analytic in the neighborhood of the point* $y^i=0$, *will be the components of a metric tensor in normal co-ordinates if, and only if, the functions* $g_{\alpha\beta,\gamma\delta}$ *satisfy* (*b*).

With the exception of the two cases treated above equations (1.5) and (1.10) furnish additional conditions over those which result by extension from the complete sets of identities (a) and (b) respectively. The determination of the exact arbitrariness of the functions $A^i{}_{jk\alpha}(y)$ or $g_{\alpha\beta,\gamma\delta}(y)$ in finite form for $n\geqq 3$ can be made on the basis of the numbers $N(n,p)$ by an extension of the above method. For this purpose we must derive a new form for the expressions $N(n,p)$.

4. *A new Form of the Expressions* $N(n,p)$. The formulas for $N(n,p)$ which determine the number of independent components $A^i{}_{jk\alpha_1\ldots\alpha}$ or $g_{\alpha\beta,\gamma_1\ldots\gamma_p}$, given in M § 1, were originally derived from equations

Affine case: $N(n,p)=nK(n,2)\,K(n,p)-nK(n,p+2)$, $\quad p\geqq 1$

Metric case: $N(n,p)=K(n,2)K(n,p)-nK(n,p+1)$, $\quad p\geqq 2$,

where $K(n,p)$ denotes the number of combinations with repetitions of n things taken p at a time.* Since in addition to the above expressions for $N(n,p)$ we will need to have the expressions for $N(n,p+r)$ before us we note that

$$(4.1)\qquad \text{Affine case:}\quad N(n,p+r)=nK(n,2)K(n,p+r)-nK(n,p+r+2)$$

$$(4.2)\qquad \text{Metric case:}\quad N(n,p+r)=K(n,2)K(n,p+r)-nK(n,p+r+1).$$

Let us now determine numbers A_i and B_i depending only on n and p such that the above expressions for $N(n,p+r)$ can be written

$$(4.3)\qquad \text{Affine case:}\quad N(n,p+r)=\sum_{i=0}^{n-1}A_iK(n-i,r)$$

$$(4.4)\qquad \text{Metric case:}\quad N(n,p+r)=\sum_{i=0}^{n-1}B_iK(n-i,r)$$

for all integer values r. To derive (4.3) and (4.4) we make use of the formula

$$K(n,p)=K(n,p-1)+K(n-1,p)$$

by repeated application of which we have

$$K(n,r+1)=K(n,r)+K(n-1,r)+\cdots+K(2,r)+1,$$

* See *Annals of Mathematics*, Vol. 28 (1927), pp. 196-236 and *ibid.*, pp. 631-688.

or

$$(4.5) \qquad K(n, r+1) = \sum_{i=0}^{n-1} K(n-i, r) = \sum_{\alpha=1}^{n} K(\alpha, r).$$

Hence

$$\begin{aligned} K(n, p+r) &= \sum_{\alpha=0}^{n-1} K(n-\alpha, p+r-1) \\ &= \sum_{\alpha=0}^{n-1} \sum_{\beta=0}^{n-\alpha-1} K(n-\alpha-\beta, p+r-2) \\ &\cdots\cdots\cdots\cdots\cdots \\ &= \sum_{\alpha=0}^{n-1} \cdots \sum_{\epsilon=0}^{n-\alpha-\cdots-\eta-1} K(n-\alpha-\cdots-\eta-\epsilon, r), \end{aligned}$$

where the indices $\alpha\beta\cdots\eta\epsilon$ are p in number. Each of the indices $\alpha\beta\cdots\eta\epsilon$ can take on values $0, 1, \cdots, n-1$, subject to the condition that the sum $\alpha+\beta+\cdots+\eta+\epsilon$ has a value between zero and $n-1$ inclusive. The number of times that the sum $\alpha+\beta+\cdots+\eta+\epsilon$ takes on the value $i\ (=0, 1, \cdots, n-1)$ is equal to the number of permutations with repetitions of the numbers $0, 1, \cdots, i$ taken p at a time such that the sum of the numbers in any permutation is equal to i; let us denote the number of times that the sum $\alpha+\beta+\cdots+\eta+\epsilon$ takes on the value $i\ (=0, 1, \cdots, n-1)$ by $S(i+1, p-1)$. More generally $S(q+1, r-1)$ will denote the number of permutations with repetitions of $0, 1, \cdots, i$ taken r at a time such that the sum of the numbers in any permutation is equal to $q (\leqq i)$. Now observe that $S(q+1, r-1)$ is also equal to the number of permutations with repetitions of $0, 1, \cdots, i$ taken $r-1$ at a time such that the sum of the numbers in any permutation is $\leqq q$ for $q \leqq i$; this is seen immediately on account of the fact that from any permutation of this latter type one of the type which go to make up the sum $S(q+1, r-1)$ can be obtained by the addition of one of the numbers $0, 1, \cdots, i$, and the fact that in this way all permutations which make up the sum $S(q+1, r-1)$ can be constructed. Hence

$$S(q+1, r-1) = \sum_{\alpha=1}^{q+1} S(\alpha, r-2)$$

for $q \leq i$. Using this formula as a recurrence formula we have

$$(4.6) \qquad S(i+1, p-1) = \sum_{\alpha=1}^{i+1} \sum_{\beta=1}^{\alpha} \cdots \sum_{\rho=1}^{\cdot} S(\rho, 1).$$

But from (4.5) an exactly similar expression can be obtained for $K(i+1,$ [illegible]

[illegible]

to $K(i+1,\, p-1)$. The above expression for $K(n, p+r)$ can therefore be written

$$(4.7) \qquad K(n, p+r) = \sum_{i=0}^{n-1} K(i+1, p-1)\, K(n-i, r).$$

Substituting this expression for $K(n, p+r)$ and a similar expression for $K(n, p+r+2)$ into (4.1) we obtain (4.3), where

$$(4.8) \qquad A_i = nK(n,2)\, K(i+1, p-1) - nK(i+1, p+1).$$

Similarly by making a substitution of the type (4.7) into (4.2) we deduce (4.4), where

$$(4.9) \qquad B_i = K(n,2)\, K(i+1, p-1) - nK(i+1, p).$$

Let us also observe that

$$(4.10) \qquad \text{Affine case:} \qquad N(n,p) = \sum_{i=0}^{n-1} A_i$$

$$(4.11) \qquad \text{Metric case:} \qquad N(n,p) = \sum_{i=0}^{n-1} B_i.$$

These equations follow from (4.8) and (4.9) when use is made of (4.5).

In the following work we shall only need (4.3) for $p=1$ and (4.4) for $p=2$, i.e.

$$(4.12) \qquad \text{Affine case:} \qquad N(n, 1+r) = \sum_{i=0}^{n-2} A_i K(n-i, r)$$

$$(4.13) \qquad \text{Metric case:} \qquad N(n, 2+r) = \sum_{i=0}^{n-2} B_i K(n-i, r).$$

For these cases A_i and B_i are given by the special formulas

$$(4.14) \qquad A_i = nK(n,2) - nK(i+1, 2) \qquad (p=1)$$

$$(4.15) \qquad B_i = K(n,2)\, K(i+1, 1) - nK(i+1, 2) \qquad (p=2).$$

It is to be noted that $A_{n-1}=0$ and $B_{n-1}=0$ is a consequence of these latter formulas.

5. *Special Case of three Variables.* Let us derive in detail the existence theorems for the functions $A^i{}_{jk\alpha}(y)$ and $g_{\alpha\beta,\gamma\delta}(y)$ for the case of $n=3$ before entering upon the more general discussion in § 6. Now divide the 24 $(=N(3,1))$ independent components $A^i{}_{jk\alpha}$ into a group G_0 consisting of 15 $(=A_0)$ components and a group G_1 consisting of 9 $(=A_1)$ components; similarly divide the 6 $(=N(3,2))$ independent components $g_{\alpha\beta,\gamma\delta}$ into a

group G_0 consisting of 3 $(= B_0)$ components and a group G_1 consisting of 3 $(= B_1)$ components. The following tables give these groups

Affine case: Group G_0				
a^i	b^i	c^i	d^i	e^i
A^i_{121}	A^i_{131}	A^i_{221}	A^i_{231}	A^i_{331}

Affine case: Group G_1		
u^i	v^i	w^i
A^i_{132}	A^i_{232}	A^i_{332}

Metric case: Group G_0		
α	β	γ
$g_{11,22}$	$g_{11,23}$	$g_{11,33}$

Metric case: Group G_1		
ζ	η	θ
$g_{21,23}$	$g_{21,33}$	$g_{22,33}$

in which the letters $a^i, b^i, \cdots, w^i$ and $\alpha, \cdots, \theta$ have been used as abbreviations for the components $A^i_{jk\alpha}$ and $g_{\alpha\beta,\gamma\delta}$. Now give the index β the value 1 and the indices $jk\alpha$ in equations (1. 6) the values (132), (232), (332) respectively as in Group G_1 for the affine case; similarly put $\epsilon = 1$ and give the indices $\alpha\beta\gamma\delta$ in (1. 12) the values (2131), (2133), (2233) as in group G_1 for the metric case. We thus arrive at the following two systems of equations

$$
\begin{aligned}
2(\partial u^i/\partial y^1) &= -(\partial d^i/\partial y^1) + 3(\partial b^i/\partial y^2) - 3(\partial a^i/\partial y^3) + \bigstar \\
6(\partial v^i/\partial y^1) &= 2(\partial u^i/\partial y^2) + 4(\partial d^i/\partial y^2) - 3(\partial c^i/\partial y^3) + \bigstar \\
3(\partial w^i/\partial y^1) &= 3(\partial e^i/\partial y^2) + 2(\partial u^i/\partial y^3) - 2(\partial d^i/\partial y^3) + \bigstar
\end{aligned} \tag{5.1}
$$

and

$$
\begin{aligned}
2(\partial \zeta/\partial y^1) &= (\partial \beta/\partial y^2) - (\partial \alpha/\partial y^3) + \bigstar \\
(\partial \eta/\partial y^1) &= (\partial \gamma/\partial y^2) - (\partial \beta/\partial y^3) + \bigstar \\
(\partial \theta/\partial y^1) &= (\partial \eta/\partial y^2) - 2(\partial \zeta/\partial y^3) + \bigstar,
\end{aligned} \tag{5.2}
$$

where the $\bigstar$ terms involve the components $C^i_{jk}, a^i, \cdots, w^i$ in (5. 1) and $h_{ij}, \partial h_{ij}/\partial y^k, \alpha, \cdots, \theta$ in (5. 2). Since these terms result from the covariant derivatives in (1. 6) and (1. 12) they are of evident form.

Let us now put

$$
a^i = A^i(y); \quad b^i = B^i(y); \quad c^i = C^i(y); \quad d^i = D^i(y); \quad e^i = E^i(y), \tag{5.3}
$$

where the $A^i, \cdots, E^i$ are arbitrary functions of the variables y^1, y^2, y^3 each of which is analytic in the neighborhood of the values $y^1 = y^2 = y^3 = 0$; also put

$$
(u^i)_{y^1=0} = P^i(y^2, y^3); \quad (v^i)_{y^1=0} = Q^i(y^2, y^3); \quad (w^i)_{y^1=0} = R^i(y^2, y^3), \tag{5.4}
$$

where the P^i, Q^i, R^i are arbitrary analytic functions in the neighborhood of the values $y^2 = y^3 = 0$. Then by evaluating (5. 1) at $y^1 = 0$ and by dif- [illegible]

[illegible]

or more differentiations with respect to the variable y^1.* In carrying out this process it is of course to be understood that $C^i_{jk} = 0$ at $y^i = 0$; also that the derivative of C^i_{jk} with respect to y^a at $y^i = 0$, i. e. $(A^i_{jka})_0$, is to be expressed in terms of the quantities $a^i, \cdots, w^i$ at $y^i = 0$; and that more generally the derivative of C^i_{jk} with respect to $y^a\, y^{\beta_1} \cdots y^{\beta_r}$ at $y^i = 0$ which we have denoted by $(A^i_{jka\beta_1 \ldots \beta_r})_0$ is determined by equations of the type $M(2.10)$. The functions (5.3) and (5.4) being given the power series expansions of all functions A^i_{jka}, namely

$$(5.5) \quad A^i_{jka} = A^i_{jka}(0) + A^i_{jka,\beta_1}(0)y^{\beta_1} + (1/2!)A^i_{jka,\beta_1\beta_2}(0)y^{\beta_1}y^{\beta_2} + \cdots,$$

is uniquely determined. As so determined the number of arbitrary quantities $A^i_{jka}(0)$ is $24\,[= N(3,\ 1)]$ and the number of arbitrary quantities $A^i_{jka,\beta_1 \ldots \beta_r}(0)$ is $15K(3, r) + 9K(2, r)$; this latter number is a consequence of the 15 arbitrary functions $A^i, \cdots, E^i$ and the 9 arbitrary functions P^i, Q^i, R^i. But

$$N(3, 1 + r) = A_0K(3, r) + A_1K(2, r)$$

from (4.12); and since $A_0 = 15$, $A_1 = 9$ by (4.14) the above number of arbitrary quantities $A^i_{jka,\beta_1 \ldots \beta_r}(0)$ is equal to $N(3, 1 + r)$. Hence the equations (1.5) can not give additional conditions on the $A^i_{jka,\beta_1 \ldots \beta_r}(0)$ over those which result from (5.1). The constant terms $A^i_{jka}(0)$ and the coefficients $A^i_{jka,\beta_1 \ldots \beta_r}(0)$ in the series (5.5) therefore satisfy all conditions required for the functions A^i_{jka} to be the components of a normal tensor in normal co-ordinates. As the series (5.5) converge † we can state the following theorem.

THEOREM D. *Let $A^i(y)$, $B^i(y)$, $C^i(y)$, $D^i(y)$, $E^i(y)$ denote* 15 *functions of the three variables y^1, y^2, y^3 and $P^i(y^2, y^3)$, $Q^i(y^2, y^3)$, $R^i(y^2, y^3)$ denote 9 functions of the two variables y^2, y^3 each function $A^i, \cdots, R^i$ being analytic in the neighborhood of the values $y^1 = y^2 = y^3 = 0$ of its arguments. Then there exists one, and only one, affine connection with components C^i_{jk} $(= C^i_{kj})$ in a system of normal co-ordinates, each function C^i_{jk} being analytic in the neighborhood of the values $y^i = 0$, such that the components $a^i, \cdots, w^i$ of the resulting normal tensor $A^i_{jka}(y)$ are*

---

* An illustration of the method of calculating these derivatives is given by E. Goursat, "Leçons sur l'Integration des équations aux dérivées partielles du premier ordre," (1921), p. 2. The system (5.1) reduces to the type considered by Goursat when the substitution (5.3) is made and the ★ terms are neglected.

† Invariantive systems of partial differential equations (see footnote on p. 227).

[illegible]

$$[illegible]\ E(y);\ [illegible]\ U^i(y);\ [illegible]\ V^i(y);\ [illegible]\ W^i(y),$$

where the U^i, V^i, W^i *are analytic functions of the variables* y^1, y^2, y^3 *in the neighborhood of the values* $y^1 = y^2 = y^3 = 0$ *such that* $U^i = P^i(y^2, y^3)$, $V^i = Q^i(y^2, y^3)$, *and* $W^i = R^i(y^2, y^3)$ *for* $y^1 = 0$.

An analogous theorem can be proved for the functions $g_{\alpha\beta,\gamma\delta}$. Here we consider that α, β, γ are arbitrary analytic functions of the three variables y^1, y^2, y^3 and that ζ, η, θ reduce to arbitrary analytic functions of the two variables y^1, y^2 for $y^1 = 0$. We assign initial values $h_{\alpha\beta} = (g_{\alpha\beta})_0 = (g_{\beta\alpha})_0$ subject to the condition that the determinant $|(g_{\alpha\beta})_0| \neq 0$, and we impose the condition $\partial h_{\alpha\beta}/\partial y^\gamma = 0$ at the point $y^i = 0$. We note that the derivative of $\partial h_{\alpha\beta}/\partial y^\gamma$ with respect to y^δ at the point $y^i = 0$, i. e. $g_{\alpha\beta,\gamma\delta}(0)$, is expressible in terms of the quantities $\alpha, \cdots, \theta$ at $y^i = 0$; also that the derivative of $\partial h_{\alpha\beta}/\partial y^\gamma$ with respect to $y^\delta, y^{\epsilon_1} \cdots y^{\epsilon_r}$ at the point $y^i = 0$, i. e., $g_{\alpha\beta,\gamma\delta\epsilon_1\ldots\epsilon_r}(0)$, is determined by equations of the type $M(2.16)$. As in the case above treated we can determine in a unique manner the successive coefficients of the power series expansions of the functions $g_{\alpha\beta,\gamma\delta}(y)$, namely

$$(5.6) \qquad g_{\alpha\beta,\gamma\delta} = g_{\alpha\beta,\gamma\delta}(0) + g_{\alpha\beta,\gamma\delta,\epsilon_1} y^{\epsilon_1} + (1/2!)\, g_{\alpha\beta,\gamma\delta,\epsilon_1\epsilon_2}(0)\, y^{\epsilon_1} y^{\epsilon_2} + \cdots.$$

Since the number of arbitrary quantities $g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}$ as above determined is $3K(3, r) + 3K(2, r)$ or $N(3, 2+r)$ by (4.13) and (4.15) the equations (1.10) can not give additional conditions on them. The formal series (5.6) therefore satisfies all conditions required for the functions $g_{\alpha\beta,\gamma\delta}$ to be the components of a metric tensor in normal co-ordinates. Since the series (5.6) converge * we can therefore state the following theorem.

THEOREM E. *Let* $a(y)$, $b(y)$, $c(y)$ *denote three functions of the three variables* y^1, y^2, y^3 *and let* $p(y^2, y^3)$, $q(y^2, y^3)$, $r(y^2, y^3)$ *denote three functions of the two variables* y^2, y^3 *and* [illegible] $a, \cdots, r$ *being analytic in the* [illegible]

$= 0$, *such that* $h_{\alpha\beta}(0) = (g_{\alpha\beta})_0$, *and such that the components* $\alpha, \cdots, \theta$ *of the resulting metric tensor* $g_{\alpha\beta,\gamma\delta}(y)$ *are*

$$\alpha = a(y)\,;\ \beta = b(y)\,;\ \gamma = c(y)\,;\quad \zeta = f(y)\,;\ \eta = g(y)\,;\ \theta = h(y),$$

where f, g, h *are analytic functions of the variables* y^1, y^2, y^3 *in the neighborhood of the values* $y^1 = y^2 = y^3 = 0$, *such that* $f = p(y^2, y^3)$, $g = q(y^2, y^3)$, *and* $h = r(y^2, y^3)$ *for* $y^1 = 0$.

In place of the arbitrary components $a^i, \cdots, e^i$ and α, β, γ in Theorems D and E certain other components can evidently be chosen as those which are completely arbitrary. It can in fact be arranged so that one of the partially arbitrary components will reduce to an arbitrary analytic function of the two variables y^2, y^3 for $y_1 = 0$, another to an arbitrary analytic function of the two variables y^1, y^3 for $y^2 = 0$, etc. From equations (5.1) and (5.2) all such possibilities can easily be deduced.

6. *General Case of* $n (\geqq 3)$ *Variables.* Equations (4.12) suggest that it is possible to divide the $N(n, 1)$ independent components $A^i{}_{jk\alpha}$ into $n - 1$ groups, namely a group G_0 comprising A_0 components each of which can be taken as a completely arbitrary function of the n variables y^i, a group G_1 comprising A_1 components each of which can be taken to reduce to an arbitrary function of the $n - 1$ variables $y^2, \cdots, y^n$ for $y^1 = 0, \cdots$, and a group G_{n-2} comprising A_{n-2} components each of which can be taken to reduce to an arbitrary function of the variables y^{n-1}, y^n for $y^1 = \cdots = y^{n-2} = 0$. Such a division of the components $A^i{}_{jk\alpha}$ into groups G_m must comprise all of the $N(n, 1)$ independent components $A^i{}_{jk\alpha}$; this is shown by (4.10). Similar remarks apply to the components $g_{\alpha\beta,\gamma\delta}$ on account of equations (4.11) and (4.13).

Let us first consider the $A^i{}_{jk\alpha}$. It is evident that the group G_0 can not be filled by selecting components $A^i{}_{jk\alpha}$ from the $N(n, 1)$ independent components $A^i{}_{jk\alpha}$ at random since they are conditioned by equations (1.6). Our problem is to show that it is possible to solve equations (1.6) for first derivatives of the independent components $A^i{}_{jk\alpha}$ such that the quantities for which we have solved will fall into $n - 2$ mutually exclusive groups; first a group consisting of first derivatives of A_1 of the independent components $A^i{}_{jk\alpha}$ with respect to y^1 (these A_1 components $A^i{}_{jk\alpha}$ will form the group G_1), second a group consisting of first derivatives of A_2 of the independent components $A^i{}_{jk\alpha}$ with respect to y^1 and y^2 (these A_2 components $A^i{}_{jk\alpha}$ will form the group G_2), $\cdots$, and finally a group consisting of first derivatives of

A_{n-2} of the independent components $A^i{}_{jk\alpha}$ with respect to y^1, $y^2, \cdots$ and y^{n-2} (these A_{n-2} components $A^i{}_{jk\alpha}$ forming the group G_{n-2}). It is possible to state a rule for the determination of the independent components $A^i{}_{jk\alpha}$ in the groups G_m $(m = 0, \cdots, n-2)$ as follows.

Rule. The group G_m $(m = 0, \cdots, n-2)$ for the components $A^i{}_{jk\alpha}$ is composed of all components that can be formed from $A^i{}_{jk\alpha}$ by taking $\alpha = m+1$; $i = 1, \cdots, n$; $j, k = 1, \cdots, n$ subject to the inequalities $j \leqq k$ and $k > m+1$. It can readily be verified that the application of this Rule gives a correct division of the components $A^i{}_{jk\alpha}$ into groups G_m. It is to be observed that the components $A^i{}_{jk\alpha}$ in a particular group G_m as given by the above Rule are independent among themselves, i. e. are not related by an equation of the form (a) and are also independent of the components $A^i{}_{jk\alpha}$ in the other groups G_m. The number of components $A^i{}_{jk\alpha}$ in the group G_m $(m = 0, \cdots, n-2)$ as determined by the above Rule is $nK(n, 2) - nK(m+1, 2)$ or A_m by (4. 14). Now let A_{lm}, where $l = 1, \cdots, A_m$; $m = 0, \cdots, n-2$, denote the A_m components $A^i{}_{jk\alpha}$ in the group G_m. On the basis of the above Rule it must therefore be possible to write equations (1. 6) in the following form in which the arrangement into groups G_m $(m = 1, \cdots, n-2)$ is also indicated:

$$
(6.1) \qquad
\begin{array}{ll}
\partial A_{l1}/\partial y^1 = \sum (\partial A_{pq}/\partial y^r) + \bigstar & (l = 1, \cdots, A_1) \\
\left\{ \begin{array}{l} \partial A_{l2}/\partial y^1 = \sum (\partial A_{pq}/\partial y^r) + \bigstar \\ \partial A_{l2}/\partial y^2 = \sum (\partial A_{pq}/\partial y^r) + \bigstar \end{array} \right. & \begin{array}{l} (l = 1, \cdots, A_2) \\ (l = 1, \cdots, A_2) \end{array} \\
\cdots\cdots\cdots\cdots & \\
\cdots\cdots\cdots\cdots & \\
\left\{ \begin{array}{l} \partial A_{l\,n-2}/\partial y^1 = \sum (\partial A_{pq}/\partial y^r) + \bigstar \\ \vdots \\ \partial A_{l\,n-2}/\partial y^{n-2} = \sum (\partial A_{pq}/\partial y^r) + \bigstar \end{array} \right. & \begin{array}{l} (l = 1, \cdots, A_{n-2}) \\ \vdots \\ (l = 1, \cdots, A_{n-2}). \end{array}
\end{array}
$$

The summations Σ in the right members of the above equations denote linear homogeneous expressions with constant coefficients in the derivatives $\partial A_{pq}/\partial y^r$. No derivative $\partial A_{pq}/\partial y^r$ which appears in the right member occurs also in the left member of any of these equations. The $\star$ terms denote polynomials in the components A and C such as result from the [illegible] in (1. 6).

It is now necessary to show that equations (1. 6) can be put into the form (6. 1). For this purpose let us first write (1. 6) as

$$(6.2)\qquad \frac{\partial A^i{}_{jk\alpha}}{\partial y^\beta} = \frac{\partial A^i{}_{jk\beta}}{\partial y^\alpha} + \frac{2}{3}\frac{\partial A^i{}_{j\beta\alpha}}{\partial y^k} + \frac{2}{3}\frac{\partial A^i{}_{k\beta\alpha}}{\partial y^j} + \frac{1}{3}\frac{\partial A^i{}_{\alpha\beta j}}{\partial y^k} + \frac{1}{3}\frac{\partial A^i{}_{\alpha\beta k}}{\partial y^j} + \star.$$

Then choose the indices $ijk\alpha\beta$ so that the derivatives in the left member of (6.2) will be equal successively to the derivatives in the left members of (6.1). We thus arrive at a set of equations

$$(6.3)\qquad \partial A_{lm}/\partial y^k = \sum (\partial A_{uv}/\partial y^w) + \bigstar.$$
$$(l = 1, \cdots, A_m,\ \ m = 1, \cdots, n-2,\ k = 1, \cdots, m).$$

For definiteness let us now consider an equation (6.2) for which $\alpha = 2$ and $\beta = 1$, i. e.

$$(6.4)\qquad \frac{\partial A^i{}_{jk2}}{\partial y^1} = \frac{\partial A^i{}_{jk1}}{\partial y^2} + \frac{2}{3}\frac{\partial A^i{}_{j12}}{\partial y^k} + \frac{2}{3}\frac{\partial A^i{}_{k12}}{\partial y^j} + \frac{1}{3}\frac{\partial A^i{}_{21j}}{\partial y^k} + \frac{1}{3}\frac{\partial A^i{}_{21k}}{\partial y^j} + \star,$$

where it is to be supposed that $j \leqq k$ and $k \geqq 3$. The component $A^i{}_{jk2}$ whose derivative stands in the left member of (6.4) therefore belongs to the group G_1. With regard to the right member of (6.4) we see that $A^i{}_{jk1}$ must belong to group G_0. If $j = 1$ the component $A^i{}_{j12}$ belongs to group G_0 since it is equal to $-2A^i{}_{121}$ by (a); similarly if $j = 2$ the component $A^i{}_{j12}$ belongs to group G_0. If $j \geqq 3$ the component $A^i{}_{j12}$ belongs to group G_1. Since $k \geqq 3$ the next component $A^i{}_{k12}$ belongs to group G_1. For $j = 1, 2$ the component $A^i{}_{21j}$ belongs to group G_0 and for $j \geqq 3$ this component is equal to $-(A^i{}_{1j2} + A^i{}_{2j1})$ of which $A^i{}_{1j2}$ belongs to group G_1 and $A^i{}_{2j1}$ belongs to group G_0. Finally $A^i{}_{21k}$ is equal to $-(A^i{}_{1k2} + A^i{}_{2k1})$ and since $k \geqq 3$ the component $A^i{}_{1k2}$ belongs to group G_1 and the component $A^i{}_{2k1}$ belongs to group G_0. As a consequence of the above it follows that the first set of equations (6.3), i. e.

$$\partial A_{l1}A_l/\partial y^1 = \sum \partial A_{uv}/\partial y^w + \bigstar \qquad (l = 1, \cdots, A_1),$$

contains in its right member derivatives of components A_{uv} which belong to groups G_0 and G_1 alone. Now consider equations (6.2) such that $\beta \leqq \mu$, $\alpha = \mu + 1$, $j \leqq k$ and $k > \mu + 1$. The component $A^i{}_{jk\alpha}$ whose derivative forms the left member of (6.2) therefore belongs to group G_μ. Let us now examine the components $A^i{}_{jk\alpha}$ whose derivatives are in the right member of (6.2) bearing in mind that $\alpha = \mu + 1$ and that j, k, β satisfy the above inequalities. The first component $A^i{}_{jk\beta}$ belongs to group $G_{\beta-1}$. The com-

[illegible] longs to group [illegible] [illegible] to group $G_{\mu-1}$. Lastly if $j < \mu$ [illegible] the component $A^i{}_{\alpha\beta}$ is to be put equal [illegible] [illegible]

to group $G_{\mu-1}$. Since k [illegible] 2 the component $A^i{}_{k\alpha\beta}$ belongs to group G. The component $A_{\alpha\beta}$ belongs to group $G_{\mu-1}$ if $j \leq \mu$ and to group G_μ if $j = \mu + 1$. If $j > \mu + 1$ the component $A^i{}_{\alpha\beta}$ is to be put equal to [illegible] where [illegible] belong to group G and [illegible] to group G. The last component $A^i{}_{\alpha\beta}$ is to be put equal to $-(A^i{}_{\beta\gamma} + A^i{}_{\gamma\beta})$ where it is seen that $A^i{}_{\rho\alpha}$ belongs to group G_μ and that $A^i{}_{k\alpha\beta}$ belongs to group $G_{\rho-1}$. This examination shows that the right members of (6. 2) in which the indices j, k, α, β are subject to the above conditions contain derivatives of components $A^i{}_{j\alpha}$ belonging to groups G_0 to G_μ inclusive. Hence in the right members of the μ-th set of equations (6. 3), i. e. the set for which $m = \mu$, there can occur only derivatives of components $A^i{}_{jk\alpha}$ which belongs to groups G_0 to G_μ inclusive.

It can be verified directly from (6. 2) that the derivative in the left member of any equation (6. 3) can occur in the right member of the same equation with a coefficient which is not equal to 1, so that the equation can be solved for the derivative in question. We can therefore consider that equations (6. 3) are such that the derivative in the left member of any equation does not occur in the right member of the same equation. Let us now suppose that a derivative $\partial A_{lm}/\partial y^k$ occurring in the left member of one of the equations (6. 3), say the pth equation, occurs also in the right member of the qth equation of (6. 3). It will be assumed that the pth equation corresponds to a value of the index m equal to M_1 and that the qth equation corresponds to a value of m equal to M_2, such that $M_1 < M_2$. Now multiply the pth equation by the coefficient of the above derivative $\partial A_{lm}/\partial y^k$ in the qth equation and combine with the qth equation by adding corresponding members of these two equations. We thus obtain an equation from which the above derivative $\partial A_{lm}/\partial y^k$ can be cancelled. Moreover the equation so obtained will contain as its left member the derivative in the left member of the qth equation and this derivative can not be cancelled [illegible]

determine the μth set of equations (6.3). From the derivatives $\partial A^i{}_{k\beta\alpha}/\partial y^j$ and $\partial A^i{}_{\alpha\beta k}/\partial y^j$ for $j \leqq \mu$ in the right members of (6.2), and from these derivatives alone, can we obtain derivatives occurring in the left members of the μth set of equations (6.3). Since $A^i{}_{\alpha\beta k} = -A^i{}_{\beta k\alpha} - A^i{}_{k\alpha\beta}$ where $A^i{}_{\beta k\alpha}$ belongs to group G_μ and $A^i{}_{k\alpha\beta}$ belongs to group G_ν where $\nu < \mu$, only the derivatives $\partial A^i{}_{k\beta\alpha}/\partial y^j$ for $j < \mu$ need be considered. We can therefore write

$$p\text{th equation:} \qquad \frac{\partial A^i{}_{jk\alpha}}{\partial y^\beta} = \frac{1}{3}\,\frac{\partial A^i{}_{\beta k\alpha}}{\partial y^j} + \cdots$$

$$q\text{th equation:} \qquad \frac{\partial A^i{}_{\beta k\alpha}}{\partial y^j} = \frac{1}{3}\,\frac{\partial A^i{}_{jk\alpha}}{\partial y^\beta} + \cdots,$$

where the dots denote quantities which do not enter into consideration. When the derivative $\partial A^i{}_{jk\alpha}/\partial y^\beta$ in the right member of the qth equation is eliminated by means of the pth equation we obtain an equation which can be solved for the derivative $\partial A^i{}_{\beta k\alpha}/\partial y^j$. This latter equation will replace the qth equation in the set (6.3). By continuing this process of combining equations (6.3) it is obviously possible to obtain equations (6.1). For the special case $n = 3$ the components $A^i{}_{jk\alpha}$ in groups G_0 and G_1 as determined by the above Rule are given in the tables in § 5; also equations (6.1) assume the particular form (5.1).

The operations performed on equations (6.3) in order to deduce the μth set of equations (6.1) involved only those equations (6.3) for which $m \leqq \mu$. Hence the above result concerning the groups of the components $A^i{}_{jk\alpha}$ whose derivatives occur in the right members of (6.3) applies to equations (6.1). This result is stated as the following lemma.

LEMMA I. *For any derivative $\partial A_{pq}/\partial y^r$ in the right members of the set of equations*

$$\partial A_{lm}/\partial y^k = \sum \partial A_{pq}/\partial y^r + \bigstar$$

$$(m = \mu, \quad 1 \leqq \mu \leqq n-2, \quad l = 1, \cdots, A_\mu, \quad k = 1, \cdots, \mu)$$

selected from (6.1) *the inequalities $q \leqq \mu$, $r > q$ are satisfied.* Denoting the derivative in the left member of any equation (6.1) by $\partial A_{lm}/\partial y^k$ so that $k \leqq m$ we can show that

$$\frac{\partial^{s+1} A_{lm}}{\partial y^k \partial y^{\alpha_1} \cdots \partial y^{\alpha_s}} = \sum \frac{\partial^{s+1} A_{pq}}{\partial y^r \partial y^{\sigma_1} \cdots \partial y^{\sigma_s}} + \sum \frac{\partial^s \bigstar}{\partial y^{\alpha_1} \cdots \partial y^{\alpha_s}}, \tag{6.5}$$

where $s = 1, 2, \cdots$; $\alpha_1, \cdots, \alpha_s = 1, \cdots, n$; $r, \sigma_1, \cdots, \sigma_s > q$, i.e. each of the indices $r, \sigma_1, \cdots, \sigma_s$ is greater than q. Each of the summations $\sum$ represents a linear homogeneous expression with constant coefficients in the

derivatives of the type indicated. It is evident that (6.5) holds for $\alpha_1, \cdots, \alpha_s = n-1, n$ since (6.5) then results directly by differentiation of (6.4). Now differentiate any equation (6.5) for $\alpha_1, \cdots, \alpha_s = n-1, n$ with respect to y^{n-2}. Then all derivatives in the right member of the resulting equation will be of the type which appears in the right member of (6.5) except derivatives of A_{n-2}. These latter derivatives, however, can be eliminated by a substitution (6.5) for $\alpha_1, \cdots, \alpha_s = n-1, n$; hence (6.5) is true for $\alpha_1 = n-2$; $\alpha_2, \cdots, \alpha_s = n-1, n$. Differentiating any equation (6.5) for $\alpha_1 = n-2$; $\alpha_2, \cdots, \alpha_s = n-1, n$ with respect to y^{n-2} we find in a similar manner that (6.5) is true for $\alpha_1, \alpha_2 = n-2$; $\alpha_3, \cdots, \alpha_s = n-1, n$; etc. Hence (6.5) holds for $\alpha_1, \cdots, \alpha_s = n-2, n-1, n$. If we now differentiate any equation (6.5) for $\alpha_1, \cdots, \alpha_s = n-2, n-1, n$ with respect to y^{n-3} and eliminate the derivatives of $A_{m\,n-3}$, $A_{m\,n-2}$ which occur, we obtain (6.5) for $\alpha_1 = n-3$; $\alpha_2, \cdots, \alpha_s = n-2, n-1, n$. Then differentiating (6.5) for $\alpha_1 = n-3$; $\alpha_2, \cdots, \alpha_s = n-2, n-1, n$, we obtain (6.5) for $\alpha_1, \alpha_2 = n-3$; $\alpha_3, \cdots, \alpha_s = n-2, n-1, n$, etc. Hence (6.5) is true for $\alpha_1, \cdots, \alpha_s = n-3, n-2, n-1, n$. Continuing this process we finally obtain (6.5) for $\alpha_1, \cdots, \alpha_s = 1, \cdots, n$, as was to be proved.

We next denote by ϕ_{lm}, where $l = 1, \cdots, A_m$; $m = 0, 1, \cdots, n-2$, an arbitrary function of the variables $y^{m+1}, y^{m+2}, \cdots, y^n$ analytic in the neighborhood of the values $y^{m+1} = \cdots = y^n = 0$. Now put $A_{l0} = \phi_{l0}$ $(l = 1, \cdots, A_0)$; also put $A_{lm} = \phi_{lm}$ $(l = 1, \cdots, A_m;\ m = 1, \cdots, n-2)$ for $y^1 = \cdots = y^m = 0$. From the functions ϕ_{lm} and equations (6.5) we can then calculate the successive coefficients $A^i{}_{jk\alpha,\beta_1\ldots\beta_r}(0)$ in the power series expansions of the functions $A^i{}_{jk\alpha}$ about the point $y^i = 0$, i.e. we can determine the series

$$(6.6) \qquad A^i{}_{jk\alpha} = A^i{}_{jk\alpha}(0) + A^i{}_{jk\alpha,\beta_1}(0)y^{\beta_1} + (1/2!)A^i{}_{jk\alpha,\beta_1\beta_2}y^{\beta_1}y^{\beta_2} + \cdots,$$

where the constant terms $A^i{}_{jk\alpha}(0)$ are given by the functions ϕ_{lm} evaluated at $y^i = 0$. In the determination of the quantities $A^i{}_{jk\alpha,\beta_1\ldots\beta_r}(0)$ the initial

conditions of integrability which would reduce the number of arbitrary quantities $A^i{}_{jka,\beta_1\ldots\beta_r}(0)$ to a number less than $N(n, 1+r)$. Equations (1.5) cannot therefore give additional conditions on the quantities $A^i{}_{jka,\beta_1\ldots\beta_r}(0)$. so that the series (6.6) if convergent will define a set of functions $A^i{}_{jka}(y)$ which will be the components of a normal tensor in a system of normal coordinates. Since it can be shown that the formal power series (6.6) converge * it follows that by Theorem A we therefore have the following theorem.

THEOREM F. *Let $\phi_{lm}(y)$, where $l = 1, \cdots, A_m$; $m = 0, 1, \cdots, n-2$, denote an arbitrary function of the variables $y^{m+1}, y^{m+2}, \cdots, y^n$ which is analytic in the neighborhood of the values $y^{m+1} = \cdots = y^n = 0$; also let A_{lm}, where $l = 1, \cdots, A_m$, denote the components $A^i{}_{jka}$ in the group G_m $(m = 0, 1, \cdots, n-2)$. Then there exists one, and only one, affine connection with components $C^i{}_{jk}(y) = C^i{}_{kj}(y)$ in a system of normal coordinates y^i, each function $C^i{}_{jk}$ being analytic in the neighborhood of the values $y^i = 0$, such that the components A_{lm} of the resulting normal tensor are*

$$A_{l0} = \phi_{l0} \qquad (l = 1, \cdots, A_0);\ \textit{and}$$
$$A_{lm} = \phi_{lm} \qquad (l = 1, \cdots, A_m;\ m = 1, \cdots, n-2)$$

for $y^1 = \cdots = y^m = 0$.

As the discussion of the functions $g_{\alpha\beta,\gamma\delta}$ is to be made in an analogous manner the special details of the argument can be omitted. The separation of the $g_{\alpha\beta,\gamma\delta}$ into groups G_m $(m = 0, \cdots, n-2)$ made on the basis of the equations

$$\frac{\partial g_{\alpha\beta,\gamma\delta}}{\partial y^\epsilon} = \frac{1}{2}\left[\frac{\partial g_{\epsilon\beta,\gamma\delta}}{\partial y^\alpha} + \frac{\partial g_{\alpha\epsilon,\gamma\delta}}{\partial y^\beta} + \frac{\partial g_{\alpha\beta,\epsilon\delta}}{\partial y^\gamma} + \frac{\partial g_{\alpha\beta,\gamma\epsilon}}{\partial y^\delta}\right] + \bigstar, \tag{6.7}$$

which are equivalent to (1.12), is given by the following rule.

Rule. The group G_m $(m = 0, \cdots, n-2)$ for the components $g_{\alpha\beta,\gamma\delta}$ is composed of all components that can be formed from $g_{\alpha\beta,\gamma\delta}$ by taking $\alpha = m+1$; $\beta, \gamma, \delta = 1, \cdots, n$ subject to the inequalities $\beta \leqq m+1$, $\beta < \gamma$, $\gamma \leqq \delta$ and $\delta > m+1$.

The components $g_{\alpha\beta,\gamma\delta}$ in the above groups G_m are to be taken as independent which can be done since they are not connected by relations of the form (b). The number of components $g_{\alpha\beta,\gamma\delta}$ in the group G_m $(m = 0, \cdots, n-2)$ is $(m+1)\,K(n, 2) - nK(m+1, 2)$ which by (4.15) is equal to B_m. Letting B_{lm}, where $l = 1, \cdots, B_m$; $m = 0, 1, \cdots, n-2$, denote

* Invariantive systems of partial differential equations (see footnote on p. 227).

[illegible]

[illegible]

$$\partial B_{l1}/\partial y^1 = \sum (\partial B_{pq}/\partial y^r) + \star \qquad (l = 1, \cdots, [illegible])$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\left\{\begin{aligned} \partial B_{l,n-2}/\partial y^1 &= \sum (\partial B_{pq}/\partial y^r) + \star \qquad (l = 1, \cdots, B_{n-2}) \\ &\vdots \\ \partial B_{l,n-2}/\partial y^{n-2} &= \sum (\partial B_{pq}/\partial y^r) + \star \qquad (l = 1, \cdots, B_{n-2}). \end{aligned}\right.$$

Equations (6.8) constitute the solution of equations (6.7) for the derivatives which appear in the left members of (6.8). For the special case $n = 3$ equations (6.8) assume the particular form (5.2). For this case also the components $g_{\alpha\beta,\gamma\delta}$ in the groups G_0 and G_1 are given in the tables in § 5. The following lemma corresponds to Lemma I.

LEMMA II. *The inequalities $q \leqq \mu$, $r > q$ are satisfied for any derivative $\partial B_{pq}/\partial y^r$ in the right members of the set of equations*

$$\partial B_{lm}/\partial y^k = \sum (\partial B_{pq}/\partial y^r) + \star$$
$$(m = \mu, \quad 1 \leqq \mu \leqq n-2, \quad l = 1, \cdots, B_\mu, \quad k = 1, \cdots, \mu)$$

selected from (6.8). Corresponding to (6.5) we now have the equations

$$\text{(6.9)} \qquad \frac{\partial^{s+1} B_{lm}}{\partial y^k \partial y^{\alpha_1} \cdots \partial y^{\alpha_s}} = \sum \frac{\partial^{s+1} B_{pq}}{\partial y^r \partial y^{\sigma_1} \cdots \partial y^{\sigma_s}} + \sum \frac{\partial^s \star}{\partial y^{\alpha_1} \cdots \partial y^{\alpha_s}},$$

where $k \leqq m$; $s = 1, 2, \cdots$; $\alpha_1, \cdots, \alpha_s = 1, \cdots, n$; and $r, \sigma_1, \cdots, \sigma_s < q$. Denoting by ψ_{lm}, where $l = 1, \cdots, B_m$; $m = 0, 1, \cdots, n-2$, an arbitrary function of the variables $y^{m+1}, \cdots, y^n$ analytic in the neighborhood of the values $y^{m+1} = \cdots = y^n = 0$, we put $B_{l0} = \psi_{l0}$ $(l = 1, \cdots, B_0)$, and for $y^1 = \cdots = y^m = 0$ we put $B_{lm} = \psi_{lm}$ $(l = 1, \cdots, B_m;\ m = 1, \cdots, n-2)$. Now assign the initial conditions $h_{\alpha\beta} = (g_{\alpha\beta})_0 = (g_{\beta\alpha})_0$ such that the determinant $|(g_{\alpha\beta})_0| \neq 0$ and $\partial h_{\alpha\beta}/\partial y^\gamma = 0$ at $y^i = 0$; also put $\partial^2 h_{\alpha\beta}/\partial y^\gamma \partial y^\delta$ at $y^i = 0$ equal to $g_{\alpha\beta,\gamma\delta}(0)$, which is determined from the given functions ψ_{lm}, and determine the higher derivatives of $h_{\alpha\beta}$ at $y^i = 0$ by equations [illegible]

[illegible]

equations (1.10) can not furnish additional conditions on the quantities $g_{\alpha\beta,\gamma\delta,\epsilon_1\ldots\epsilon_r}(0)$.

It can be shown that the series (6.10) converge;* hence by theorem B we have the following theorem.

THEOREM G. *Let ψ_{lm}, where $l=1,\cdots,B_m$, $m=0, 1,\cdots,n-2$, denote an arbitrary function of the variables $y^{m+1}, y^{m+2},\cdots,y^n$ which is analytic in the neighborhood of the values $y^{m+1}=\cdots=y^n=0$; let $(g_{\alpha\beta})_0=(g_{\beta\alpha})_0$, where $\alpha, \beta=1,\cdots,n$, be arbitrary constants such that the determinant $|(g_{\alpha\beta})_0|\neq 0$; also let B_{lm}, where $l=1,\cdots,B_m$, denote the components $g_{\alpha\beta,\gamma\delta}$ in the group G_m $(m=0, 1,\cdots,n-2)$. Then there exists one, and only one, fundamental tensor with components $h_{\alpha\beta}(y)=h_{\beta\alpha}(y)$ in a system of normal co-ordinates y^i, each function $h_{\alpha\beta}(y)$ being analytic in the neighborhood of the values $y^i=0$, such that $h_{\alpha\beta}(0)=(g_{\alpha\beta})_0$, and the components B_{lm} of the resulting metric tensor are*

$$B_{l0}=\psi_{l0} \qquad (l=1,\cdots,B_0);\ \textit{and}$$
$$B_{lm}=\psi_{lm} \qquad (l=1,\cdots,B_m;\ m=1,\cdots,n-2)$$

for $y^i=\cdots=y^m=0$.

PRINCETON, N. J.

* Invariantive systems of partial differential equations (see footnote on p. 227).

Relations Between the Critical Points of a Real Analytic Functions of N Independent Variables.*

By Arthur B. Brown.†

1. *Introduction. A critical point of a function is one where all its first partial derivatives are zero.* Birkhoff was the first to state a relation between the critical points of a function of n variables.‡ Morse § went further, obtaining a complete set of relations in the important (and general) case of non-degenerate critical points; that is, where the Hessian of the function is not zero at any critical point. We shall find it convenient to borrow certain results from his paper. Morse's results are described in the second part of the present paper.

In this paper we investigate the case of the general isolated critical point, dropping the restriction used by Morse, but requiring the function to be analytic. Relations similar to Morse's are obtained. They may be described roughly by saying that each general isolated critical point has the effect of a whole group of non-degenerate critical points.

I. Preliminary Theorems of Analysis Situs.

2. *Notations and definitions.* We shall use the nomenclature of analysis situs as defined by Alexander,¶ substituting the terms *cycle* and *Betti number* for *closed chain* and *connectivity number* respectively. Notations not found in Alexander's article will be as defined in Veblen's Colloquium Lectures on Analysis Situs.

In the entire paper all geometric terms and operations such as Betti number, addition of chains, bounding, may be taken either modulo 2 or absolute. For definiteness we shall take the ordinary (absolute) case, but it is understood that the work would go through as well in the modulo 2 case.

* Presented to the American Mathematical Society, June 20, 1929, in slightly different form.

† National Research Fellow.

‡ For references on critical points see the author's paper "Relations between the [illegible]," *Annals of Mathematics*, Vol. 31 (1930).

§ Marston Morse, "Relations between the Critical Points of a Real Function of n Independent Variables," *Transactions of the American Mathematical Society*, Vol. 27 (1925), pp. 345-396.

¶ [illegible], Vol. 2[illegible] ([illegible]), pp. [illegible].

3. *Relations between nested complexes.*

THEOREM 1. *Let C and K be given complexes, where K contains C in the point-set sense. Let a deformation D exist which keeps points of K on K, and carries K into a complex on C. Then the Betti numbers of C are at least as great as those of K.*

Proof. Let A be a complete set of non-bounding i-cycles of C, where i is any non-negative integer. The theorem will be proved if we can show that any i-cycle of K is dependent* on cycles of A.

Let E be any i-cycle of K. Let E_1 be the transform of E by D. Then E_1 may be a singular i-cycle.

Since A is a complete set of non-bounding i-cycles of C, we know that there exists an integer $m \neq 0$ and a combination A_2 of cycles of A such that

$$mE_1 + A_2 \sim 0 \quad \text{on } C, \text{ hence also on } K.$$

Now

$$mE - mE_1 \sim 0 \quad \text{on } K,$$

as follows from the properties of a deformation. By addition of the two homologies we obtain the following.

$$mE + A_2 \sim 0 \quad \text{on } K.$$

Hence the theorem is proved.

THEOREM 2. *Let C and K be given complexes, where K contains C in the point-set sense. Let a deformation D exist which keeps points of K on K, keeps points of C on C, and carries K into a complex on C. Then C and K have the same Betti numbers.*†

Proof. Let A be a complete set of non-bounding i-cycles of C, where i is any non-negative integer. Now the hypotheses of Theorem 1 are satisfied, and in proving Theorem 1 we showed that any i-cycle of K is dependent on cycles of A. If we can show that there is no homology on K among the cycles of A, it will follow that A is a complete set of non-bounding i-cycles of K, and Theorem 2 will be proved.

Let A_1 be any linear combination of cycles of A, with integral coefficients, not all zero. Let A_2 be the transform of A_1 by D. Since the transforms of A_1 at all intermediate stages of D lie on C, the following homology is valid.

---

* Dependence shall always refer to homologies.

† It may be seen from the proofs that Theorems 1 and 2 apply to any geometrical configurations for which Betti numbers, topologically defined, exist.

[illegible]

[illegible]

Now if A_1 bounded any complex, say E, of K, then A_2 would [illegible] [illegible] the following homology.

$$A_2 \sim 0 \quad \text{on } C.$$

Adding to the last homology, we would have the following.

$$A_1 \sim 0 \quad \text{on } C.$$

But that is impossible, since A is a complete set of non-bounding i-cycles of C. Hence there are no homologies on K among the cycles of A, and Theorem 2 is true.

4. *Relations between the Betti numbers of overlapping complexes.* The following theorem will be used later in the paper to obtain relations between the critical points.

THEOREM 3. *Let A and B be complexes of dimensionalities m and n respectively, $m \leqq n$. Let the points common to A and B form a complex C. Let the set of all points that belong to either A or B, or both, form a complex D which can be so divided into cells that at the same time A, B and C receive proper divisions into cells.*

Then integers α_i, $i = 0, 1, \cdots, n$; and β_j, $j = 0, 1, \cdots, n+1$; exist, all positive or zero, with $\beta_0 = \beta_{n+1} = 0$, such that the following relations hold: *

$$R_i(C) = \beta_{i+1} + \alpha_i, \tag{4.1}$$

$$R_i(D) - R_i(A) - R_i(B) = \beta_i - \alpha_i, \quad i = 0, 1, \cdots, n. \tag{4.2}$$

Proof. The proof will require a number of steps, which we shall designate as lemmas. We begin by introducing notation for certain sets of cycles.

Let A^i and B^i be maximum sets of i-cycles of A and B respectively, which are independent, on A and B respectively, of cycles of C.

Let [illegible] and [illegible] be maximum sets of i-cycles of C which bound [illegible]

[illegible]

The number of cycles in any set will be denoted by replacing the large letter by a small letter. Thus $C_1{}^i$ contains $c_1{}^i$ cycles.

LEMMA 1. $R_i(A) = a^i + c_1{}^i + c_b{}^i$.

Proof. The sets A^i, $C_1{}^i$, $C_b{}^i$ are composed of cycles of A. That they contain a complete set of non-bounding i-cycles of A is easily verified from the definitions above. If we supposed a homology existed among them on A, consideration of these definitions will show successively that it could involve no cycles of A^i, none of $C_1{}^i$, and none of $C_b{}^i$. Hence there is no such homology, and the lemma is true.

LEMMA 2. $R_i(B) = b^i + c_1{}^i + c_a{}^i$.

Proof. This follows from Lemma 1 because of the symmetry between A and B.

LEMMA 3. *The cycles $C_1{}^i$, $C_a{}^i$, $C_b{}^i$, $C_d{}^i$ are independent on C.*

Proof. If there were any homology among them, it could not involve any cycles of $C_1{}^i$, since the latter are independent on D, and all the others are bounding on D. Then it could involve no cycles of $C_a{}^i$, since they are independent on B. For a similar reason it could involve no cycles of $C_b{}^i$. Hence it must be a homology, on C, among the cycles of $C_d{}^i$, which is contrary to the definition of that set. Therefore there can be no such homology, and the lemma is proved.

LEMMA 4. $R_i(C) = c_1{}^i + c_a{}^i + c_b{}^i + c_d{}^i$.

Proof. In view of Lemma 3 it will be sufficient to show that if E^i is any i-cycle of C, then it is dependent on cycles of the sets mentioned in Lemma 3.

From the definition of $C_1{}^i$ it follows that there exists an integer $m \neq 0$, and a (possibly empty) combination $E_1{}^i$ of cycles of $C_1{}^i$ such that

$$(4.3) \qquad mE^i + E_1{}^i \sim 0 \quad \text{on } D.$$

Let P^i be the boundary of the part in A of an $(i+1)$-chain K^{i+1} corresponding to the last homology. Then

$$(4.4) \qquad F^i = mE^i + E_1{}^i - P^i \sim 0 \quad \text{on } B.$$

That is because F^i is the boundary of

$$L^{i+1} = (K^{i+1} \text{ minus the part of } K^{i+1} \text{ in } A).$$

Now either there is a non-zero multiple of F^i that bounds in A, or there is not. In the first case F^i must be dependent, on C, on cycles of C_d^i, and in the second it must be dependent, on A, on cycles of C_b^i. That follows from the definitions of C_d^i and C_b^i. It is therefore correct to say, regardless of which case is at hand, that an integer $q \neq 0$ and a (possibly empty) combination E_b^i of cycles of C_b^i and C_d^i, exist such that

$$qF^i + E_b^i \sim 0 \quad \text{on } A. \tag{4.5}$$

By similar proof we can show that there exists an integer $r \neq 0$ and a combination E_a^i of cycles of C_a^i and C_d^i, such that

$$rP^i + E_a^i \sim 0 \quad \text{on } B. \tag{4.6}$$

In proving this we would use, in place of (4. 4), the homology

$$P^i \sim 0 \quad \text{on } A. \tag{4.7}$$

From (4. 4) and (4. 7) we conclude that (4. 5) holds also on B, and (4. 6) holds also on A. Thus each holds both on A and on B. From the definition of C_d^i we conclude that there exist integers s and t, neither zero, and (possibly empty) combinations H_1^i and H_2^i of cycles of C_d^i, such that

$$sqF^i + sE_b^i + H_1^i \sim 0 \quad \text{on } C,$$
$$trP^i + tE_a^i + H_2^i \sim 0 \quad \text{on } C.$$

Let us multiply the first of these homologies by tr, the second by sq, and add. One term of the result is

$$trsq(F^i + P^i),$$

which, on comparison with (4. 4), can be written as

$$trsq(mE^i + E_1^i).$$

The other terms are combinations of cycles of C_a^i, C_b^i and C_d^i. It follows that E^i is dependent, on C, on cycles of C_a^i, C_b^i, C_1^i and C_d^i; and Lemma 4 is proved.

Before stating the next lemma we shall introduce a new set of cycles. Given any cycle of C_d^{i-1}, it bounds both in A and in B, by definition of C_d^{i-1}. If it bounds D_a^i in A and D_b^i in B, then $D_a^i - D_b^i$ has no boundary, hence is a cycle. Thus for each cycle of C_d^{i-1} can be found a corresponding cycle of

Let D^i be a set of i-cycles of D obtained in this way. We note that c_d^{i-1} is the number of cycles in D^i.

LEMMA 5. *The cycles A^i, B^i, C_1^i, D^i are independent on D.*

Proof. Suppose we have a homology on D among the four sets of cycles. We shall show that it can involve no cycles.

Corresponding to the homology is a relation of bounding,

$$E^{i+1} \to P_a^i + P_b^i + P_1^i + P_d^i. \tag{4.8}$$

Here the last four symbols denote the combinations of cycles of A^i, B^i, C_1^i, D^i, respectively, involved.

Let E_a^{i+1} be the part of E^{i+1} in A, and let F^i be the boundary of E_a^{i+1}. Let

$$E_b^{i+1} = E^{i+1} - E_a^{i+1},$$

and let H^i be the boundary of E_b^{i+1}.

Let K^{i-1} be the combination of cycles of C_d^{i-1} corresponding to the combination P_d^i of cycles of D^i. Let L_a^i and L_b^i be the chains of A and respectively, bounded by K^{i-1}, such that

$$P_d^i = L_a^i - L_b^i.$$

Now $(F^i - P_a^i - L_a^i)$ is an i-chain of C bounded by $(-K^{i-1})$, for the first two terms in the first parentheses represent cycles. Since the cycles of C_d^{i-1} are independent on C, it follows that K^{i-1} must be a combination of them with all zero coefficients. Therefore P_d^i contains no cycles.

Since F^i and H^i are the boundaries of E_a^{i+1} and E_b^{i+1} respectively, the following are correct relations:

$$E_a^{i+1} \to P_a^i + (F^i - P_a^i), \tag{4.9}$$

$$E_b^{i+1} \to P_b^i + (H^i - P_b^i). \tag{4.10}$$

Now E_a^{i+1} is a chain of A, and $(F^i - P_a^i)$ is a set of cycles of C. It follows from (4.9) that P_a^i is homologous on A to cycles of C. Since P_a^i is a combination of cycles of A^i, such a homology is impossible unless P_a^i is empty. Hence P_a^i contains no cycles, and by similar proof, using (4.10), P_b^i is empty.

Therefore (4.8) tells us that P_1^i is bounding on D. Since the cycles of C_1^i are independent on D, it follows that P_1^i is empty. Hence the given homology involves a combination of the cycles of A^i, B^i, C_1^i, D^i with all coefficients zero. We conclude that Lemma 5 is true.

[illegible]

LEMMA 6. [illegible]

Proof. In view of Lemma 5 it will be sufficient to show that if E^i is any i-cycle of D, it must [illegible].

Let E_a^i be the part of E^i in A, and $E_b^i = E^i - E_a^i$. Let F^{i-1} be the boundary of E_a^i. Then F^{i-1} is a cycle of C, and bounds $(-E_b^i)$ as well.

From Lemma 1 we conclude that F^{i-1} is dependent, on C, on cycles of C_1^{i-1}, C_a^{i-1}, C_b^{i-1} and C_d^{i-1}. Now C_1^{i-1} cannot enter in the dependence, for if it did the combination of cycles of C_1^{i-1} in question would be bounding on D, inasmuch as each of the four other sets bounds on D. Neither can C_a^{i-1}, since C_b^{i-1}, C_d^{i-1} and F^{i-1} all bound on B. For a similar reason, C_b^{i-1} can contribute nothing to the dependence. Hence a bounding relation of the following sort exists.

$$L^i \rightarrow mF^{i-1} + H^{i-1}.$$

Here L^i is a chain of C, m is a non-zero integer, and H^{i-1} is a combination of cycles of C_d^{i-1}.

Let K_a^i and K_b^i be chains of A and B respectively, bounded by H^{i-1}. That they exist follows from the definition of C_d^i. Then

$$\begin{gathered}(mE_a^i - L^i + K_a^i) \text{ is a cycle of } A, \text{ and}\\ (mE_b^i + L^i - K_b^i) \text{ is a cycle of } B.\end{gathered}$$

According to Lemmas 1 and 2 these sums are, then, dependent on cycles of A^i, B^i, C_1^i, C_a^i, C_b^i. Since the last two sets contain cycles that bound on D, it follows that on D the two sums in parentheses are dependent on cycles A^i, B^i and C_1^i. Therefore the sum of the two parentheses, namely

$$mE^i + (K_a^i - K_b^i)$$

is dependent on cycles of A^i, B^i and C_1^i. But $K_a^i - K_b^i$ is a combination of cycles of D^i, as follows from the definitions of K_a^i, K_b^i and D^i. We conclude that E^i is dependent on cycles of A^i, B^i, C_1^i and D^i, as was to be proved.

[illegible] Theorem 3. L[illegible] 3 [illegible]

[illegible]

no $(n+1)$-cycles, it follows that there are no cycles in C_d^n, and therefore $c_d^n = \beta_{n+1} = 0$. This completes the proof.

COROLLARY 1. $\alpha_i \leqq R_i(A) + R_i(B)$. $\beta_i \leqq R_i(D)$.

Proof. These inequalities can be verified by a glance at Lemmas 1 and 2, and the definitions of α_i and β_i.

COROLLARY 2. *Let A_i, B_i, C_i, D_i denote the i-th Betti numbers of A, B, C, D respectively. Then the following relations hold:*

$$(C_0 + D_0) \geqq (A_0 + B_0) \geqq D_0,$$
$$(C_0 + D_0) - (C_1 + D_1) \leqq (A_0 + B_0) - (A_1 + B_1) \leqq (C_0 + D_0) - D_1,$$
$$(C_0 + D_0) - (C_1 + D_1) + (C_2 + D_2) \geqq (A_0 + B_0) - (A_1 + B_1) + (A_2 + B_2)$$
$$\geqq (C_0 + D_0) - (C_1 + D_1) + D_2,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$(C_0 + D_0) - \cdots + (-1)^n (C_n + D_n)$$
$$= (A_0 + B_0) - \cdots + (-1)^n (A_n + B_n).$$

Proof. The relations can be verified by substitution from (4.1) and (4.2) and using the relations $\beta_0 = \beta_{n+1} = 0$.

5. *Result of adding a zero-cell.*

LEMMA 7. *Let K be a given complex and P a point not in K, nor connected with it. If P is added to K, the only change in Betti numbers is an increase of unity in R_0.*

Proof. The lemma is obvious and is only stated for convenience in reference.

LEMMA 8. *Let P be a point of an ordinary n-space, and K a generalized complex in a finite part of the space, containing all the points of the space neighboring P except the point P itself. If P is added to K, the only change in Betti numbers (topologically defined), is a decrease of one in R_{n-1}.*

Proof. The addition does not change R_0, the number of parts. It does not change $R_1, R_2, \cdots$, or R_{n-2}, for any cycles of dimensionalities $1, 2, \cdots, n-2$ which pass through P can be deformed into cycles not passing through P; and any chains of dimensionalities $2, 3, \cdots, n-1$ bounded by such cycles, and passing through P, can likewise be deformed so as not to pass through P. Neither does the addition change R_n, since neither of the complexes in question fills out all n-space.

But it does reduce R_{n-1} by unity, as we show now. Let S be a small $(n-1)$-sphere with center at P. Then S, originally non-bounding, becomes bounding. Suppose U were an $(n-1)$-cycle independent of S, which became bounding when P is added to K. Then U could be deformed so as to be outside S, and since the bounded chain must contain a certain non-zero multiple, say m, of the complex consisting of S and its interior, it follows that U would be homologous on K to mS, a contradiction. (Each chain mentioned is supposed to have a definite orientation.) It follows that R_{n-1} is reduced by only one, and the lemma is proved.

II. Critical Points.

6. *Hypotheses.* The following is familiar terminology. A function of class $C^{(k)}$ is one continuous and possessing continuous partial derivatives of orders $1, 2, \cdots, k$. A locus is a regular m-spread in n-space if near any of its points it is given by expressing $(n-m)$ of the co-ordinates as functions, of class C', of the other m co-ordinates as independent variables. It is of class $C^{(k)}$ if the functions are of class $C^{(k)}$.

The hypotheses for part II are given in the next paragraph.

A closed region R of a real euclidean n-space is given, bounded by regular $(n-1)$-spreads of class $C^{(3)}$. A single-valued function f is given, real and analytic over R. Its inner normal derivative, at any boundary point, is negative. Its critical points are isolated.

Morse proved * that under these hypotheses it is possible to redefine f near the boundary so that it assumes a constant value on the boundary, greater than the value at any interior point. No new critical points are introduced by the change, and the new function is of class C''. Hereafter f will denote the new function.

7. *Morse's results.* The essential differences in the hypotheses used by Morse are as follows. The function f is not required to be analytic, but to be of class C''. The Hessian of f must not vanish at any critical point.

The critical points are then proved to be finite in number. They are divided into $n+1$ classes, defined as follows. Let

$$a_{ij} = f_{x_i x_j},$$

---

* [illegible]

evaluated at a given critical point. Then the form* $a_{ij}x_ix_j$ can be carried by a real, non-singular linear transformation into the form

$$(7.1) \qquad -y_1^2 - \cdots - y_k^2 + y_{k+1}^2 + \cdots + y_n^2.$$

The critical point is said to be of type k, where $k=0$ if all the signs in (7.1) are positive.

Let R_i be the i-th Betti number, modulo 2, of R, and M_i the number of critical points of type i, $i=0, 1, \cdots, n$. Then a set of n inequalities and one equation were obtained by Morse, in form identical with those given in Theorem 10 of this paper.

The same relations hold if the Betti numbers are taken absolute instead of modulo 2.

The corresponding results in this paper are stated in Theorem 10.

8. *Some properties of the critical points.* Morse showed by a short proof that the hypothesis of isolated critical points leads to the following conclusion.†

THEOREM 4. *The number of critical points is finite.*

A slight extension of a result demonstrated by Morse ‡ yields the following lemma.

LEMMA 9. *Let P be any point not a critical point, where f is analytic. The loci $f=$ constant neighboring P are regular analytic $(n-1)$-spreads.§ Their orthogonal trajectories are regular and analytic, form a field, and can be represented by the equations*

$$x_i = x_i(x_1^0, \cdots, x_n^0, t), \qquad i=1, 2, \cdots, n,$$

where $(x_1^0, \cdots, x_n^0)$ is the point of intersection of the trajectory with the locus $f=$ constant through P, the x_i are analytic functions, and $t=f$ is the parameter on the trajectory.

Hereafter we shall refer to the orthogonal trajectories to the manifolds $f=$ constant simply as orthogonal trajectories.

---

* We shall use the notation in which repetition of a subscript denotes summation from 1 to n.

† Morse, *loc. cit.*, page 348.

‡ Morse, *loc. cit.*, pp. 355-358.

§ All loci mentioned are real, unless the contrary is stated.

THEOREM 5. *If a point is isolated on a locus $f = c$, then it is a point either of relative minimum or of relative maximum, of f.*

Proof. If every neighborhood of the point, say P, contained points satisfying $f < c$ and points satisfying $f > c$, we could join a point of the first set to a point of the second set by a curve near P but not passing through P. The curve would then contain a point at which $f = c$, contrary to the hypothesis that there are no such points near P. Hence the theorem is true.

Any critical point not a point of minimum or of maximum shall be called a critical point of intermediate type.

LEMMA 10. *Let the origin of co-ordinates be a critical point of intermediate type at which $f = c$. Then in a sufficiently small neighborhood of the origin the function*

$$H \equiv (x_i f_{x_j} - x_j f_{x_i})(x_i f_{x_j} - x_j f_{x_i})$$

does not vanish at any real point of the locus $f = c$ other than the origin itself.

Geometrically, this lemma states that the normal to the locus $f = c$ does not pass through the origin.

Proof of Lemma 10.*

The points in question are the real points at which the following equations are satisfied:

$$x_i f_{x_j} - x_j f_{x_i} = 0, \qquad i, j = 1, 2, \cdots, n; \tag{8.1}$$

$$f - c = 0. \tag{8.2}$$

The points, real and complex, not the origin, satisfying these relations and neighboring the origin, lie on a finite number of configurations g of the following nature.† A configuration of degree s, $0 < s < n$, is defined when $n - s$ of the variables are expressed as single-valued functions of the other variables, say $x_1, x_2, \cdots, x_s$, on a configuration defined by the vanishing of a pseudopolynomial for which $x_1, x_2, \cdots, x_s$ are the independent variables. These functions are analytic at points where the discriminant of the pseudopolynomial is not zero (ordinary points).

Each pseudopolynomial may be taken irreducible. Let us consider one

* The outline of this proof is due to Morse.

† Osgood, *Lehrbuch der Funktionentheorie*, Vol. II, Zweite Auflage, page 132. We shall refer to this volume as Osgood II. We shall refer to Vol. I of the *Lehrbuch*, Fünfte Auflage, as Osgood I.

of them that is irreducible. Its discriminant cannot vanish identically.* Hence the points at which it does vanish lie on sub-configurations of degrees not exceeding $s-1$, defined by the vanishing of the discriminant, a function of $x_1, \cdots, x_s$. Therefore if the configuration does not contain any real ordinary points, all its real points are confined to the configurations of lower dimensionalities which result when we set the discriminant equal to zero.

It will, then, be sufficient to consider configurations containing real ordinary points. We shall show that on each of these the function

$$r^2 = x_i x_i$$

is constant. Now if a configuration, say g_1, is of degree s_1, then in the neighborhood of any ordinary point on it, r^2 can be expressed as an analytic function of s_1 of the variables $x_1, x_2, \cdots, x_n$ as independent variables. It is possible to continue analytically from any ordinary point to any second ordinary point of g_1.† Therefore, if we can show that r^2 is constant in the neighborhood of a single ordinary point of g_1, it must equal the same constant at all ordinary points. Since r^2 is a continuous function, and there are ordinary points neighboring any point not itself an ordinary point, it will then follow that r^2 equals the same constant over the entire configuration g_1.

Let P be a real ordinary point (not the origin) on g_1. Let Q be any nearby point of g_1. Let t denote an arbitrary one of the independent variables for the configuration g_1. Since (8. 2) is satisfied, we can differentiate at Q, to obtain the following relation:

$$f_{x_i} \partial x_i / \partial t = 0. \tag{8. 3}$$

Since P is a real point not the origin, and Q is near P, it follows that Q is neither a critical point nor the origin. Hence the functions

$$x_j x_j \quad \text{and} \quad f_{xj} f_{xj}$$

are not zero at Q. Therefore, since equations (8. 1) are satisfied on g_1, the following relations hold at Q:

$$f_{x_i} = \pm \left[(f_{xj} f_{xj})^{1/2} / (x_j x_j)^{1/2} \right] \cdot x_i. \tag{8. 4}$$

Here the sign is the same for all n values of i.

Let us substitute now from (8. 4) in (8. 3), to obtain the following relation:

$$\left[(f_{xj} f_{xj})^{1/2} / (x_j x_j)^{1/2} \right] \cdot \left[x_i \cdot (\partial x_i / \partial t) \right] = 0. \tag{8. 5}$$

---

* Osgood II, page 108, 3 Satz.

† Osgood II, page 109.

Since the first factor is not zero, we conclude that the last factor is zero. Hence

$$\partial(x_i x_i)/\partial t = 0$$

at Q. Therefore all the first partial derivatives of r^2 with respect to the independent variables are zero at all points Q neighboring P. Consequently r^2 is constant in the neighborhood of P, on g_1. Therefore r^2 has the same value over the entire configuration g_1.

Hence r^2 has only a finite number of values for all real points satisfying (8.1) and (8.2) simultaneously. Therefore any sufficiently small neighborhood of the origin contains none of these points. The proof of Lemma 10 is now complete.

9. *The trajectories* τ. We shall obtain a set of regular analytic trajectories which form a field in the neighborhood of any real point at which

$$H \equiv (x_i f_{x_j} - x_j f_{x_i})(x_i f_{x_j} - x_j f_{x_i}) \neq 0. \tag{9.1}$$

Let us consider the following system of differential equations,

$$dx_i/dt = \rho[x_i(f_{x_j} f_{x_j}) - f_{x_i}(x_j f_{x_j})], \tag{9.2}$$

where

$$\rho = [2(x_i x_i)^{1/2}/H(x_1, \cdots, x_n)]. \tag{9.3}$$

The brackets in (9.2) were obtained as direction components of the intersection of the tangent $(n-1)$-plane to the locus $f = \text{constant}$ at a point P, with the 2-plane passing through the origin and the normal at P to the tangent $(n-1)$-plane.

LEMMA 11. *In the neighborhood of any point at which* (9.1) *is satisfied, there exists a set of regular analytic trajectories* τ, *defined by* (9.2) *and* (9.3), *which form a field. The function* f *is constant along any one of these trajectories, and the parameter* t *can be taken equal to* $r = (x_i x_i)^{1/2}$.

Proof. The right hand members of (9.2) are analytic. If the right hand members of any two of these equations are zero, then we can show very easily that a corresponding term of the form in (9.1) is zero. Since [illegible]

[illegible]

[illegible] $= 0$.

Let us consider the following:

$$\frac{dr}{dt}=\frac{x_i\dfrac{dx_i}{dt}}{r}=\frac{\rho[(x_ix_i)(f_{x_j}f_{x_j})-(x_if_{x_i})(x_jf_{x_j})]}{r}$$

$$=\frac{\rho(x_if_{x_j}-x_jf_{x_i})(x_if_{x_j}-x_jf_{x_i})}{2r}=1,$$

as follows from (9.3). Hence t may be taken equal to r. This completes the proof.

10. *The locus $f=$ constant in the neighborhood of a critical point of intermediate type.*

THEOREM 6. *Let P be a critical point of intermediate type at which $f=c$. Let S be an $(n-1)$-sphere with center at P, so small that* (9.1) *is satisfied at all points of the locus $f=c$ within or on S. Then the real intersection of S with the locus $f=c$ consists of a finite number of regular anialytic $(n-2)$-dimensional manifolds.*

Proof. Let B denote the set of intersection points in question. The points of B are given by the real simultaneous solutions of the equations

(10.1) $$\phi\equiv f-c=0,$$
(10.2) $$\psi\equiv(x_ix_i)-d^2=0,$$

where d is the radius of the sphere S. The Jacobians of these functions are

$$[D(\phi,\psi)/D(x_i,x_j)]=2(x_jf_{x_i}-x_if_{x_j}).$$

By hypothesis (9.1) is satisfied. Therefore not all these Jacobians are zero at any point of B. Hence* the points of B neighboring a fixed point of B are given by expressing two of the variables as real analytic functions of the other $n-2$. Consequently any connected part of B is a regular analytic $(n-2)$-spread. Since ϕ and ψ are continuous, it is a closed set of points. It can be shown to be a manifold.† The number of these manifolds is very easily shown to be finite.‡ Thus Theorem 6 is proved.

THEOREM 7. *Let P be a critical point of intermediate type, at which*

---

* Osgood I, page 70, and II, page 14.

† The hypotheses at hand are sufficient to ensure that the locus is a complex in the sense of analysis situs. Similar results will be needed in section 11. The necessary proofs have been given by B. L. van der Waerden. For reference and proofs see Lefschetz's *Colloquium Lectures on Topology*, Chapter 8.

‡ Cf. Morse, *loc. cit.*, page 348.

$f = c$. Let S be an $(n-1)$-sphere with center at P, so small that (9.1) *is satisfied at all the points, say A, of the locus $f = c$ within or on S. Let B be the intersection of S with the locus $f = c$.*

Then A can be put in one-to-one continuous correspondence with the set of all points on straight line segments joining P to points of B.

Proof. Let a correspondence be set up in which a given line segment, say L, is made to correspond to the trajectory τ on A that cuts B in the same point as L. Corresponding points shall be equi-distant from P. By use of Lemma 11 this correspondence is easily proved to be continuous.

11. *Differences in Betti numbers of the loci $f \leqq c - e$ and $f \leqq c + e$.* We shall begin by introducing notations for certain complexes that we shall use.

Any value assumed by f at a critical point is called a critical value.

Let $e > 0$ and c be constants such that c is the only critical value in the interval $c - e \leqq f \leqq c + e$. Let a set of $(n-1)$-spheres S be taken, one about each critical popint of intermediate type on $f = c$ as center, small enough to satisfy the hypotheses of Theorem 7.

Let A be the complex of all points satisfying $f \leqq c$ within or on the spheres S.

Let B be the complex of all points satisfying $f \leqq c$ except those interior to the spheres S.

Let B' be the (generalized) complex obtained by removing from B the points on $f = c$ at which f has maxima and minima.

Let C be the complex of points on the spheres S satisfying $f \leqq c$.

Let D be the complex of all points satisfying $f \leqq c$.

Let D' be the complex obtained by removing from D all the critical points at which $f = c$.

Let E and F be the complexes satisfying $f \leqq c + e$ and $f \leqq c - e$ respectively.

Let H be the complex of all points satisfying $f = c$ within or on the spheres S.

LEMMA 12. *The complexes D and E have the same Betti numbers.*

Proof. Let e_1 be taken satisfying $0 < e_1 \leqq e$ and so small that the complex of points satisfying

$$c \leqq f \leqq c + e_1$$

can be deformed onto the locus $f = c$, through points satisfying

$$c \leqq f \leqq c + e.\text{*}$$

Next we define a deformation in two steps. In the first step the points satisfying

$$c + e_1 < f \leqq c + e$$

move along their orthogonal trajectories so that f decreases at a constant rate; stopping when they reach the locus

$$f = c + e_1.$$

The second step is the deformation mentioned in defining e_1.

The deformation described in these two steps shows, by Theorem 2, that D and E have the same Betti numbers, as was to be proved.

LEMMA 13. *The complexes F and D' have the same Betti numbers.*

Proof. We make a deformation of D' along orthogonal trajectories, which moves only those of its points satisfying $c - e < f \leqq c$. This deformation may be defined in a manner similar to the first one used in the proof of Lemma 12, so as to carry these points onto the locus $f = c - e$. The deformation is possible, because there is a unique orthogonal trajectory through each point that is moved, leading to the locus $f = c - e$. We may now apply Theorem 2, and find that Lemma 13 is true.†

LEMMA 14. *The complexes B' and D' have the same Betti numbers.*

Proof. We shall make a deformation satisfying the hypotheses of Theorem 2, and carrying the points that are in D' but not in B', onto the spheres S. It will be sufficient to describe the deformation within a single one of the spheres S. Let P be the corresponding critical point.

Consider an $(n-1)$-sphere S' with center at P, starting with zero radius and then expanding at a constant rate till it coincides with S. The deformation is then defined as follows.

Any given point Q remains fixed till S' reaches that point. Thereafter Q travels on S' towards S along a fixed straight line through P, that is, so

---

* The property of a complex here implied is proved in Lefschetz's Colloquium Lectures on "Topology."

† The idea of using D', obtained from D by the removal of the critical points, was suggested to me by Professor Lefschetz. It introduced a great simplification in my original treatment, and also has enabled me to state the final results in slightly simpler form.

long as it does not reach H. If Q reaches H it then travels along a trajectory τ (see Theorem 7) on H for such time as the points on the (moving) radial line through P and Q near Q but on the side away from P, are points at which $f > c$. (But Q still remains on S'.) As soon as the case arises that these points satisfy $f \leqq c$, then Q begins again to move along a fixed line through P.

If it happens that the radial line becomes tangent to H, then the trajectory τ through the point in question has the same direction as the radial line. Therefore the motion is continuous, and constitutes a deformation.

The deformation is seen to satisfy the hypotheses of Theorem 2. Therefore D' and B' have the same Betti numbers, and the lemma is proved.

LEMMA 15. *The differences in Betti numbers of the loci $f \leqq c - e$ and $f \leqq c + e$ are the same as the corresponding differences for B' and D.*

Proof. The result follows from Lemmas 12, 13 and 14, and the definitions of E and F.

LEMMA 16. *The Betti numbers of A are $R_0 = N$, $R_1 = R_2 = \cdots = R_n = 0$; where N is the number of critical points of intermediate type at which $f = c$.*

Proof. This is proved by making a deformation carrying A into the corresponding set of N critical points, and then applying Theorem 2. The deformation is defined in a manner similar to that used in the proof of Lemma 14, using contracting spheres instead of expanding spheres. Because of the similarity to the proof of Lemma 14, it is unnecessary to give any further details.

Let P be any critical point of intermediate type, and C_p the complex of points on the sphere S about P as center, at which $f \leqq c$. *The i-th type number, $M_i(P)$, of the given critical point, is defined by*

$$M_i(P) = R_{i-1}(C_p), \qquad (i = 2, 3, \cdots, n-1);$$
$$M_1(P) = R_0(C_p) - 1.$$

Thus each critical point of intermediate type has $(n-1)$ type numbers.

[illegible]

Let C_p'' be the region obtained by projecting C_p' radially on the sphere S. We now make a stereographic projection of S onto a tangent $(n-1)$-plane L, taking the pole at a point of S not in C_p or C_p''. We shall use the same notation for the projections of C_p and C_p'' as for the original regions.

Now C_p is an $(n-1)$-region in L, bounded by regular analytic $(n-2)$-spreads. (See Theorem 6.) We can deform the part of C_p near the boundary a short distance toward the interior along the normals to the boundary, till the deform is interior to C_p''. From Theorem 1 we conclude that the Betti numbers of C_p'' are at least as great as those of C_p. By symmetry it follows that the Betti numbers of C_p are at least as great as those of C_p''. Therefore the two have the same Betti numbers, and consequently the same is true of C_p and C_p' as originally defined.

We have proved that if S is one sphere which may be used, then spheres of radii sufficiently near that of S define complexes C_p having the same Betti numbers. It follows that all suitable spheres S give complexes C_p with the same Betti numbers. For if any two are given, we can show easily by use of the Heine-Borel theorem and the above result that we can pass from one to the other with no change in Betti numbers. Therefore Theorem 8 is proved.

THEOREM 9. *Let c be a critical value and $e > 0$ a suitably chosen small constant. Let ΔM_0 and ΔM_n be the numbers of points of minimum and maximum, respectively, at which $f = c$. Let ΔM_i, $i = 1, 2, \cdots, n-1$, be the sum of the i-th type numbers of all the critical points of intermediate type at which $f = c$. Let E and F be the complexes satisfying $f \leqq c + e$ and $f \leqq c - e$ respectively, and $\Delta R_i = R_i(E) - R_i(F)$. Then the following relations are valid:*

$$\Delta M_0 \geqq \Delta R_0,$$
$$\Delta(M_0 - M_1) \leqq \Delta(R_0 - R_1),$$
$$\Delta(M_0 - M_1 + M_2) \geqq \Delta(R_0 - R_1 + R_2),$$
$$\cdots\cdots\cdots\cdots$$
$$\Delta[M_0 - M_1 + \cdots + (-1)^n M_n] = \Delta[R_0 - R_1 + \cdots + (-1)^n R_n].$$

Proof. We shall evaluate the right hand members of the above relations. According to Lemma 15 the changes in Betti numbers are those obtained when the points of minimum and maximum are added to B', and then the complex A is added (except for C, which is already part of B').

Now the minima are not connected to B', and the maxima have all their neighboring points belonging to B'. It follows from Lemmas 7 and 8 that the addition of the minima and maxima contributes M_0 to ΔR_0 and $(-M_n)$ to ΔR_{n-1}. The addition of the maxima and minima converts B' into B.

Next we add the complex A (except for the part of its boundary already in B). The resulting complex is D. The points common to A and B form the complex C. Here we can apply the second corollary to Theorem 3, at the end of section 4. The A, B, C and D of Theorem 3 will be the complexes of the same names defined at the beginning of this section. We begin by giving certain formulas and substitutions to be used in the application of the corollary.

In view of the result just obtained, we have the following formulas for the ΔR_i:

$$\begin{aligned}
&\Delta R_0 = \Delta M_0 + D_0 - B_0, \\
&\Delta R_i = D_i - B_i, \qquad i = 1, 2, \cdots, n-2, n. \\
&\Delta R_{n-1} = D_{n-1} - B_{n-1} - \Delta M_n.
\end{aligned}$$

We find from Lemma 16 that $A_0 = N$, $A_1 = A_2 = \cdots = A_n = 0$. Here N denotes the number of critical points of intermediate type at which $f = c$.

From the definition of type number we conclude that

$$\begin{aligned}
&C_0 = \Delta M_1 + N, \\
&C_i = \Delta M_{i+1}, \qquad i = 1, 2, \cdots, n-2.
\end{aligned}$$

Since C is a sum of parts of $(n-1)$-spheres, but contains no complete $(n-1)$-spheres, we know that $C_{n-1} = C_n = 0$.

The n-th Betti numbers of all the complexes are zero, since none of them fills out n-space.

We are now ready to use the right hand set of relations in the corollary, but we replace the first relation by the following:

$$A_0 + B_0 \geqq D_0 + N.$$

To prove this relation we observe that A contains N parts, each of which is connected to B. That makes the inequality obvious, since the zeroth Betti number of a complex equals the number of paprts of the complex.

After making the replacement just mentioned, we solve the right hand relations of the corollary for values of the right hand members of the relations stated in Theorem 9. By using the facts stated above, we find that the result is the set of relations forming the conclusion of the theorem. This completes the proof.

12. *Relations between the critical points.* Morse proved[*] that if there is no critical value in the interval $a \leq f \leq b$, $(a < b)$, then the com-

[*] Morse, *loc. cit.*, pp. 358-360, and 396.

plexes defined by the relations $f \leqq a$ and $f \leqq b$ have the same Betti numbers. Since the Betti numbers of an empty complex are all zero, it follows that relations involving the Betti numbers of the region R over which f is defined may be obtained by adding the relations of Theorem 9 for all critical values. Thus we get the following theorem.

THEOREM 10. *Let M_0 and M_n be the numbers of critical points of minimum and maximum, respectively, of f in R. Let M_i be the sum of the i-th type numbers of all the critical points of intermediate type in R. Let R_i denote the i-th Betti number of R. Then the following relations hold.*

$$M_0 \geqq R_0,$$
$$M_0 - M_1 \leqq R_0 - R_1,$$
$$M_0 - M_1 + M_2 \geqq R_0 - R_1 + R_2,$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$M_0 - M_1 + \cdots + (-1)^n M_n = R_0 - R_1 + \cdots + (-1)^n R_n.$$

HARVARD UNIVERSITY,
PRINCETON UNIVERSITY.

Numbers of Representations in Certain Quinary Quadratic Forms.

By E. T. Bell.

---

1. 1. *Introduction.* A general problem, indicated in §§ 1. 4, 1. 5, concerning enumerations of representations in quadratic forms is illustrated by complete enumerations for four forms in 5 variables. Several partial enumerations for similar forms are determined incidentally while discussing the four forms; these results will be omitted here. These apparently are the first instances of enumerations for quinary quadratic forms other than a sum of 5 squares. The paper also illustrates several simple, practicable methods which I have applied successfully to forms in 2, 3, 4, 5, $\cdots$, 15 variables. Selections from the numerous theorems thus obtained will be published later.

1. 2. *Notation.* Without further reference, α is an integer $\geqq 0$; n is an integer > 0; m is an odd integer > 0; $\mu, \mu_1, \mu_2 \cdots$ are odd integers $\gtrless 0$; $x, y, x, w, u, v, v_1, v_2, \cdots$ are integers $\gtreqless 0$. Restrictions consistent with any of the foregoing will always be indicated, as, for example, $m \equiv 3 \bmod 4$; such restrictions will apply only to the equation before which they are written. The number of distinct one-rowed matrices $(x, y, z, \cdots)$ such that $n = f(x, y, z, \cdots)$ is the number of representations of n in $f(x, y, z, \cdots)$, and is denoted by $N[n = f(x, y, z, \cdots)]$ or, when there can be no confusion, by $N[n = f]$. We define $\phi_\alpha(n)$ by

$$\phi_\alpha(n) \equiv N[n = x^2 + y^2 + z^2 + w^2 + 2^\alpha u^2].$$

Hence $\phi_0(n)$, which will be written $\phi(n)$, is the number of representations of n as a sum of 5 squares. The functions $\phi_\alpha(n)$, $\alpha > 0$, are fundamental for the work mentioned in § 1. 1, end.

1. 3. *Simple and Compound Enumerations.* The general and, as yet, unsolved problem in § 1 can be precisely formulated. It will be simpler, however, for the present, merely to describe the nature of the problem with precision sufficient to cover the forms actually treated in this paper, which will serve as an introduction to a more general statement.

Let f be a quadratic form such that $N[n = f]$ is finite. All $N[n = f]$ fall into two mutually exclusive classes: all $N[n = f]$ that are polynomials in the real divisors of cn alone, where c is an integer not depending upon n;

all other $N[n=f]$. Members of the first class will be called *simple;* of the second, *compound.*

The classification of compound enumerations can be carried out in several sufficient (but usually unsatisfactory) ways. For example, if the generating function of $N[n=f]$ is a modular invariant, this invariant may be expressed in terms of the members of a fundamental set, and the enumerative functions generated by the members of any such set may be taken as the elementary functions of which $N[n=f]$ is a function. This particular classification throws no light on the arithmetical properties of the $N[n=f]$ concerned.

1.4. *The Function* $\phi(n)$. It is known that $N[n=f]$ is simple when f is a sum of 2, 4, 6 or 8 squares, and compound when n is a sum of $2r (r>4)$ squares, or when n is a sum of $3, 5, 7, \cdots$ squares. Further, $\phi(n)$ is not rationally expressible in terms of the number of representations of n as a sum of 3 squares, or a sum of $2r (r>4)$ squares. Hence if a particular $N[n=f]$ is determined in the form

$$a_1\phi(b_1n)+a_2\phi(b_2n)+\cdots+a_s\phi(a_sn),$$

where $s, b_1, \cdots, b_s$ are integers >0 independent of n, and $a_1, \cdots, a_s$ are rational numbers independent of n, it is futile to seek any essentially simpler expression for $N[n=f]$.

We therefore regard $\phi(n)$ as a new elementary enumerative function, and seek those forms f such that

$$N[n=f]=c\chi(an)+a_1\phi(b_1n)+\cdots+a_s\phi(b_sn)$$

where χ is simple, and $a>0$, $c \gtreqless 0$ are independent of n.

Similar statements may be made with respect to $\phi_\alpha(n)$, $\alpha>0$, or any $N[n=f]$ which can be shown to be compound.

1.5. *Degeneration.* Let $N[n=f]$ be compound. For n suitably restricted, the compound $N[n=f]$ may degenerate to a simple function. Thus, if f is a sum of 10 squares $N[n=f]$ is compound, but $N[4\alpha+3=f]$ is simple. Again, if f is a sum of 3 or of 5 squares, it was shown by Stieltjes and Hurwitz that $N[n^2=f]$ is simple, and a similar result has recently been proved for 7 squares by Dr. Gordon Pall, who has also generalized the theorems.*

* In passing, it may be pointed out that it has not yet been proved for these $N[n=f]$ that no integers a, b exist such that $N[an+b=f]$ is simple. That is, it is not known whether there exist for $3, 5, 7, \cdots$ squares analogues of the theorems for 10 and 12 squares. This indicates the undeveloped state of the theory.

The like is not true for 9 squares; the phenomenon reappears in one case of 11 squares, beyond which nothing concerning this point is known.

The foregoing remarks suggest the problem of determining the polynomial $P(n)$ of lowest degree with integral coefficients such that, when f is given and $N[n=f]$ is compound, $N[P(n)=f]$ is simple. For the forms discussed in this paper $P(n)$ is n^2.

It is interesting to observe the evident analogies between § 1. 4, § 1. 5 and classical problems of the integral calculus.

2. 1. *Complete Enumerations for Four Special Quinaries.* Before indicating the proofs, we summarize the results for the forms $f_1, \cdots, f_4$ where

$$\begin{aligned} f_4 &\equiv x^2 + 4y^2 + 4z^2 + 4w^2 + 4u^2, \\ f_3 &\equiv x^2 + y^2 + 4z^2 + 4w^2 + 4u^2, \\ f_2 &\equiv x^2 + y^2 + z^2 + 4w^2 + 4u^2, \\ f_1 &\equiv x^2 + y^2 + z^2 + w^2 + 4u^2. \end{aligned}$$

For f_4, we shall prove the following:

$$\begin{array}{ll} m \equiv 3 \bmod 4: & N[m=f_4]=0; \\ m \equiv 1 \bmod 8: & 5N[m=f_4]=\phi(m); \\ m \equiv 5 \bmod 8: & 7N[m=f_4]=\phi(m); \\ & N[2m=f_4]=0; \\ & N[4n=f_4]=\phi(n). \end{array}$$

For f_3,

$$\begin{array}{ll} m \equiv 3 \bmod 4: & N[m=f_3]=0; \\ m \equiv 1 \bmod 8: & 5N[m=f_3]=2\phi(m); \\ m \equiv 5 \bmod 8: & 7N[m=f_3]=2\phi(m); \\ & 10N[2m=f_3]=\phi(2m); \\ & N[4n=f_3]=\phi(n). \end{array}$$

For f_2,

$$\begin{array}{ll} m \equiv 3 \bmod 4: & 10N[m=f_2]=\phi(m); \\ m \equiv 1 \bmod 8: & 5N[m=f_2]=3\phi(m); \end{array}$$

[illegible]

[illegible]

For f_1,

$$
\begin{array}{ll}
m \equiv 3 \bmod 4: & 5N[m = f_1] = 2\phi(m); \\
m \equiv 3 \bmod 4: & 35N[2^{2a+2}m = f_1] = 3(2^{3a+4} + 5)\phi(m); \\
m \equiv 1 \bmod 8: & 5N[m = f_1] = 4\phi(m); \\
m \equiv 1 \bmod 8: & 35N[2^{2a+2}m = f_1] = (3 \cdot 2^{3a+5} - 5)\phi(m); \\
m \equiv 5 \bmod 8: & 7N[m = f_1] = 4\phi(m); \\
m \equiv 5 \bmod 8: & 49N[2^{2a+2}m = f_1] = 3(2^{3a+5} + 3)\phi(m); \\
 & 35N[2^{2a+1}m = f_1] = 3(2^{3a+1} + 5)\phi(2m).
\end{array}
$$

2.2. *Degeneracies.* For the following statement of Hurwitz' theorem on 5 squares, we refer to a former paper, where full references are given.* Denote by $\zeta_r(n)$ the sum of the r^{th} powers of all the divisors of n, and define $H(m)$ by

$$H(m) \equiv \Pi[\zeta_3(p^a) - p\zeta_3(p^{a-1})], \quad H(1) = 1,$$

where $m = \Pi p^a = p^a q^b \cdots r^c$ is the resolution of m into powers of distinct primes $p, q, \cdots, r$. Then Hurwitz' theorem may be written

$$\phi(2^{2a}m^2) = 10\zeta_3(2^a)H(m).$$

All the degeneracies for $f_4, \cdots, f_1$ obtainable by application of the foregoing are read off from § 2.1. Observe that in § 2.1 only the cases $m \equiv 1 \bmod 8$ and $N[4n = f]$ need be considered, since an odd square is $\equiv 1 \bmod 8$, and an even square is $\equiv 0$ or $4 \bmod 8$. Hence

$$
\begin{array}{ll}
N[m^2 = f_4] = 2H(m), & 7N[2^{2a+2}m^2 = f_4] = 10(2^{3a+3} - 1)H(m); \\
N[m^2 = f_3] = 4H(m), & 7N[2^{2a+2}m^2 = f_3] = 10(2^{3a+3} - 1)H(m); \\
N[m^2 = f_2] = 6H(m), & 7N[2^{2a+2}m^2 = f_2] = 10(2^{3a+3} - 1)H(m); \\
N[m^2 = f_1] = 8H(m), & 7N[2^{2a+2}m^2 = f_1] = 2(3.2^{3a+5} - 5)H(m).
\end{array}
$$

Since every square is of one or other of the forms m^2, $2^{2a+2}m^2$, degeneration of $N[n = f]$ when n is a square is complete for each of the forms $f_4, \cdots, f_1$.

2.3. *Reduction Formulas.* Most of the statements in § 2.1 can be disposed of by observing the necessary and sufficient congruential forms modulo 4 of the numbers representable in a particular f. Passing these for the moment, we consider certain reduction formulas. Such seem to play an essential part in complete determinations of all but the simplest $N[n = f]$.

The forms $f_4, \cdots, f_1$ are in 5 variables with coefficients powers of 2. Hence it is suggested that we examine those theta identities of degree 5

---

* *Transactions of the American Mathematical Society*, Vol. 26 (1924), p. 448.

arising from transformations of order a power of 2. Similarly for forms in any number of variables at least one of whose coefficients is divisible by a prime p (other than 2). The forms $f_1, \cdots, f_1$ are too simple, however, to exhibit the power of this naturally suggested method, and we shall use a more direct means. The special simplification for $f_4, \cdots, f_1$ enters through the identity $\vartheta_0^4 + \vartheta_2^4 = \vartheta_3^4$, which is only one degree lower than those required. The notation is as usual

$$\vartheta_a \equiv \vartheta_a(q), \quad \vartheta_3(q) = \vartheta_0(-q) = \Sigma q^{\nu^2}, \quad \vartheta_2(q^4) = \Sigma q^{\mu^2}.$$

If the stated identity be multiplied in turn by ϑ_0, ϑ_2 ϑ_3 throughout, the three resulting identities when interpreted arithmetically give practically all of the required information for the forms considered in this paper. We need attend only to one aspect of the identity

$$\vartheta_0^4\vartheta_3 + \vartheta_2^4\vartheta_3 = \vartheta_3^5$$

as the rest of the analysis is of the same nature. Similarly, for the identities deduced from these by the transformations of orders 2, 4, 8.

Comparison of coefficients in the given identity yields

$$\begin{aligned}[n = \nu_1^2 + \nu_2^2 + \nu_3^2 + \nu_4^2 + \nu_5^2 \,:\, \Sigma(-1)^{\nu_1+\nu_2+\nu_3+\nu_4}] \\ + N[4n = \mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2 + 4\nu_5^2] = \phi(n),\end{aligned}$$

where $[n = f : \Sigma \text{ etc.}]$ denotes the indicated sum taken over all sets of solutions of $n = f$. Since precisely 0, 1, 2, 3 or 4 of $\nu_1, \nu_2, \nu_3, \nu_4$ may be odd, the last is equivalent to

$$\begin{aligned}
& N[n = \nu_1^2 + 4\nu_2^2 + 4\nu_3^2 + 4\nu_4^2 + 4\nu_5^2] \\
+ & \,6N[n = \mu_1^2 + \mu_2^2 + 4\nu_3^2 + 4\nu_4^2 + 4\nu_5^2] \\
+ & \;\;N[n = \mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2 + 4\nu_5^2] \\
- & \,4N[n = \mu_1^2 + 4\nu_2^2 + 4\nu_3^2 + 4\nu_4^2 + \nu_5^2] \\
- & \,4N[n = \mu_1^2 + \mu_2^2 + \mu_3^2 + 4\nu_4^2 + \nu_5^2] \\
+ & \;\;N[4n = \mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2 + 4\nu_5^2] = \phi(n),
\end{aligned}$$

in some lines of which obviously permissible changes of suffixes have been made. Change n into $4n$ and attend to the necessary congruential possibilities. Then

Now in $4n = f_1$, necessarily

$$x^2 + y^2 + z^2 + w^2 \equiv 0 \bmod 4,$$

and therefore only precisely 0 or 4 of x, y, z, w may be odd. Thus

$$N[4n = f_1] = \phi(n) + N[4n = \mu_1^2 + \mu_2^2 + \mu_3^2 + \mu_4^2 + 4\nu_5^2].$$

Hence if we write

$$\psi(\alpha) \equiv N[2^\alpha m = f_1], \quad \lambda(\alpha) \equiv \phi(2^\alpha m),$$

the difference equation is equivalent to

$$\psi(\alpha + 4) - 3\psi(\alpha + 2) = 2\lambda(\alpha + 2) - 4\lambda(\alpha).$$

Suppressing the simple algebra, we write down the solution. Let $\{x\} = 0$ or the greatest integer $\overline{\gtrless} x$, according as $x \overline{\gtrless} 0$ or $x > 0$. Then the solution is

$$\psi(\alpha + 4) - 3^{\{\alpha/2\}+1} \psi(\alpha - 2\{\alpha/2\} + 2)$$
$$= 2\lambda(\alpha + 2) - 4 \cdot 3^{\{\alpha/2\}} \lambda(\alpha - 2\{\alpha/2\}) + 2 \sum_{j=0}^{\{\alpha/2\}-1} 3^j \lambda(\alpha - 2j),$$

the summation being absent when $\alpha < 2$.

In general, it is difficult or impossible to proceed beyond this stage. Hence it may be pointed out that the solution implied for the number of representations of $2^{\alpha+4}m$ in f_1 is fairly satisfactory as it depends only upon the like for $4m$ or $8m$, according as α is even or odd, and a sum of functions ϕ whose arguments depend only upon α (not upon m). However, in this case the solution can be completed by means of the following very special cases of Dr. Pall's theorems *:

$$\begin{array}{ll}
m \equiv 3 \bmod 4: & 7\lambda(2\alpha) = (2^{3\alpha+2} + 3)\phi(m); \\
m \equiv 1 \bmod 8: & 7\lambda(2\alpha) = (2^{3\alpha+3} - 1)\phi(m); \\
m \equiv 5 \bmod 8: & 49\lambda(2\alpha) = (5.2^{3\alpha+3} + 9)\phi(m); \\
 & 7\lambda(2\alpha + 1) = (2^{3\alpha+2} + 3)\phi(2m).
\end{array}$$

For the complete reduction, we shall need also the following, whose origin will be indicated presently:

$$\begin{array}{ll}
m \equiv 3 \bmod 4: & 5\psi(2) = 9\phi(m); \\
m \equiv 1 \bmod 8: & 5\psi(2) = 13\phi(m); \\
m \equiv 5 \bmod 8: & 7\psi(2) = 15\phi(m).
\end{array}$$

* To be published in the *Journal of the London Mathematical Society.*

If the foregoing be applied to the solution above for $\psi(\alpha+4)$, after separation of the cases α even, α odd, all of the summary in 2. 1 for $N[n=f_1]$, except the results for $2^\alpha m$ $(\alpha=0,\ 3)$, is found at once. Passing these for the moment, all is as stated for f_1.

2. 4. *Remaining Cases.* In $2n=f_4$, x must be even, and $2m=f_4$ is impossible. In $m=f_4$, necessarily $m\equiv 1 \bmod 4$. Now in $m=x^2+y^2+z^2+w^2+u^2$, only precisely 5, 3 or 1 of x, y, z, w, u may be odd. Hence

$$\begin{aligned}\phi(m)=&N[m=\mu_1{}^2+\mu_2{}^2+\mu_3{}^2+\mu_4{}^2+\mu_5{}^2]\\&+10N[m=\mu_1{}^2+\mu_2{}^2+\mu_3{}^2+4\nu_4{}^2+4\nu_5{}^2]\\&+\ \ 5N[m=\mu_1{}^2+4\nu_2{}^2+4\nu_3{}^2+4\nu_4{}^2+4\nu_5{}^2],\end{aligned}$$

the coefficients 1, 10, 5 being supplied since the odd variables in the indicated fixed positions in the representations as a sum of 5 squares may be chosen from the available 5 in these numbers of ways. If $m\equiv 3$ mod 4 only the second N is present, and hence $10N[m=f_2]=\phi(m)$ if $m\equiv 3 \bmod 4$. If $m\equiv 1 \bmod 8$, only the third N is present, so that $5N[m=f_4]=\phi(m)$ when $m\equiv 1 \bmod 8$. Finally, if $m\equiv 5 \bmod 8$, only the first and third N's are relevant. This case can be disposed of in several ways. To illustrate one which (amplified) applies to forms in 3, 4, 5, $\cdots$, 15 variables, we cite a result which is written down immediately from two formulas in a previous paper*:

$$m\equiv 5 \bmod 8: \qquad 7N[m=\mu_1{}^2+\mu_2{}^2+\mu_3{}^2+\mu_4{}^2+\mu_5{}^2]=2\phi(m),$$

which will often be found useful. From this and the above expression for $\phi(m)$, we find the result stated in § 2. 1 for $N[m=f_4]$ when $m\equiv 5 \bmod 8$, which completes the discussion for f_4.

In $m=f_3$, precisely one of x, y must be even. Hence $N[m=f_3]=2N[m=f_4]$. Since in $4n=f_3$, both x and y must be even, $N[4n=f_3]=\phi(n)$. Again, since $2m\equiv 2 \bmod 4$, precisely 2 of the variables in $2m=x^2+y^2+z^2+w^2+u^2$ must be odd. Hence $10N[2m=f_3]=\phi(2m)$ as stated.

In $2n=f_2$, either precisely 0 or precisely 2 of x, y, z may be odd, and these exhaust the possibilities. If precisely 2 of x, y, z are odd, $n=m$. Hence $N[2m=f_2]=3N[2m=f_3]$. Clearly, $N[4n=f_2]=\phi(n)$. If $m=f_2$, then necessarily either precisely 1 or precisely 3 of x, y, z are odd. If the former, $m\equiv 1 \bmod 4$; if the latter, $m\equiv 3 \bmod 4$, and in particular

$$m\equiv 1 \bmod 8: \qquad N[m=f_2]=3N[m=f_1].$$

If $m\equiv 3 \bmod 4$, only precisely 3 of the variables in the representation of m

* *American Journal of Mathematics*, Vol. 42 (1920), pp. 177-178.

as a sum of 5 squares are odd. Thus $10N[m=f_2]=\phi(m)$ when $m\equiv 3 \bmod 4$. Finally, by an argument from congruences, as already frequently used, we find $m\equiv 1 \bmod 4$: $N[m=f_4]+N[m=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+\mu_5^2]=\phi(m)$; whence, by a similar argument, we get

$$\begin{aligned}
m&\equiv 1 \bmod 4:\ 5N[m=f_4]=\phi(m)-N[m=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+\mu_5^2];\\
m&\equiv 1 \bmod 4:\ 5N[m=f_3]=2\phi(m)-2N[m=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+\mu_5^2];\\
m&\equiv 1 \bmod 4:\ 5N[m=f_2]=3\phi(m)-3N[m=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+\mu_5^2];\\
m&\equiv 1 \bmod 4:\ 5N[m=f_1]=4\phi(m)-4N[m=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+\mu_5^2].
\end{aligned}$$

Using the value determined above for the N on the right, we find from the third of these the remaining case of f_2, and from the fourth, the stated values of $N[m=f_1]$ for $m\equiv 1 \bmod 4$.

The value of $N[m=f_1]$ when $m\equiv 3 \bmod 4$ can be obtained in several ways. For brevity, we cite

$$m\equiv 3 \bmod 4:\quad \phi(m)=20\Sigma\zeta_1(m-4\nu^2)\qquad (\nu=0,\pm 1,\pm 2,\cdots)$$

from the former paper. Hence, by the theorem on 4 squares, the result follows.

There remains only $N[8m=f_1]$, which was deferred in § 2. 3. We shall prove the following in a moment:

$$\begin{aligned}
&2N[2n=\mu_1^2+\nu_2^2+\nu_3^2+\nu_4^2+\nu_5^2]\\
&\qquad\qquad=N[8n=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+4\mu_5^2];\\
&2N[m=x^2+y^2+z^2+w^2+2u^2]\\
&\qquad\qquad=N[2m=x^2+y^2+z^2+w^2+u^2;\ x \text{ odd}];\\
&2N[2m=f_3]=N[m=x^2+y^2+z^2+w^2+2u^2].
\end{aligned}$$

In the first, take $n=m$, and recall in the last that $10N[2m=f_3]=\phi(2m)$. Then

$$\begin{aligned}
&5N[\ m=x^2+y^2+z^2+w^2+2u^2]=\phi(2m);\\
&5N[2m=x^2+y^2+z^2+w^2+u^2;\ x \text{ odd}]=2\phi(2m);\\
&5N[8m=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+\mu_5^2]=4\phi(2m).
\end{aligned}$$

Again, in $8m=f_1$, only precisely 0 or 4 of x, y, z, w may be odd. Hence

$$N[8m=f_1]=\phi(2m)+N[8m=\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2+4\mu_5^2],$$

as is easily seen from the residues mod 16. Hence from the last of the preceding, we have the value of $N[8m=f_1]$ as stated in 2. 1.

The first of the assumed theorems follows at once from

$$N_5(8n,4)=80\Sigma\zeta_1(2n-\mu^2),$$

quoted from the previous paper, *loc. cit.*, where $N_r(n,s)$ denotes the number of representations of n as a sum of r squares precisely s of which are odd. As a similar argument has already been used, we omit further details.

To vary the method of proof slightly for the second and third assumed theorems, and to bring out a historical point of some interest, we note that these theorems (with many more) are implicit in a result stated by Gauss in 1808.* To translate Gauss' formula to the present context, consider all resolutions of n into a pair of positive divisors t, τ of which τ is odd, $n = t\tau$, and write

$$\psi(n) = \Sigma(-1)^t t,$$

the Σ extending to all such t. Let $\epsilon(n) = 0$ or 1 according as n is not or is a square. Then the equivalent of Gauss' result is easily found to be the identity

$$\Sigma\psi(n - \nu^2) = -n\epsilon(n) \quad (\nu = 0, \pm 1, \pm 2, \cdots).$$

Recall that $24\zeta_1'(n)$, where $\zeta_1'(n)$ is the sum of all the odd divisors of n, is the number of representations of $2n$ as a sum of 4 squares, and observe that

$$\psi(m) = -\zeta_1'(m), \quad \psi(2n) = 2(-1)^n\psi(n).$$

Then, from the above identity, it follows immediately that

$$N[2n = x^2 + y^2 + z^2 + w^2 + u^2;\, x \text{ odd}] = 16n\epsilon(2n) + 8\Sigma\psi(2n - 4\nu^2),$$
$$N[m = f_1] = 8m\epsilon(m) + 8\Sigma\psi(m - \mu^2),$$

the last Σ referring to $\mu = \pm 1, \pm 3, \pm 5 \cdots$. In the first, take $n = m$, $n = 2m$, and reduce the resulting identities by the above. Then

$$N[2m = x^2 + y^2 + z^2 + w^2 + u^2;\, x \text{ odd}] = 2N[m = x^2 + y^2 + z^2 + w^2 + 2u^2],$$
$$N[4m = x^2 + y^2 + z^2 + w^2 + u^2;\, x \text{ odd}] = 8N[m = f_1].$$

From the first, by the argument from congruence as already often used, we find

$$N[m = x^2 + y^2 + z^2 + w^2 + 2u^2] = 2N[2m = x^2 + y^2 + 4z^2 + 4w^2 + 4u^2].$$

A similar argument on the second gives, on starting from the representation of $4m$ as a sum of 5 squares,

$$10N[m = f_1] = \phi(4m) - \phi(m);$$

whence, applying the reduction formulas for ϕ in § 2. 3, we have $N[m = f_1]$ as stated in § 2. 1.

If the detailed [illegible] ... with but little further labor.

On Groups of Motion in Related Spaces.*

By M. S. KNEBELMAN.†

1. The object of this brief note is to prove two theorems pertaining to the groups of motion of two Riemann spaces which are in geodesic correspondence or are conformal to each other.

Let V_n be an n-dimensional space whose fundamental tensor has the components $g_{ij}^{(x)}$ in the coördinate system x. If $\xi^i(x)$ are the contravariant components of an infinitesimal motion of the space into itself, this vector must satisfy Killing's equations ‡

$$\xi_{i,j} + \xi_{j,i} = 0, \tag{1.1}$$

where ξ_i $(\equiv g_{ih}\xi^h)$ are the covariant components of the motion and a subscript preceded by a comma denotes covariant differentiation with respect to the g's. The number of linearly independent solutions of the differential equations (1.1) can not exceed $n(n+1)/2$, spaces of constant Riemannian curvature being the only ones admitting a group of motions of this number of parameters.

We now suppose that another space $\bar{V}_n$, with fundamental tensor whose components in the coördinate system x are $\bar{g}_{ij}(x)$, is in geodesic correspondence with V_n. Then, as is well known,

$$\overline{\left\{ {i \atop jk} \right\}} = \left\{ {i \atop jk} \right\} + \delta_j^i\phi_{,k} + \delta_k^i\phi_{,j}, \tag{1.2}$$

where $\left\{ {i \atop jk} \right\}$ are the Christoffel symbols of the second kind formed of the g's, while $\overline{\left\{ {i \atop jk} \right\}}$ are the corresponding symbols formed of the $\bar{g}$'s; ϕ is a scalar which by contraction of (1.2) is found to be

$$\phi = [1/2(n+1)] \log (\bar{g}/g), \qquad g = | g_{ij} | \tag{1.3}$$

and δ_j^i is the Kronecker delta $(=1$ if $i=j$; $=0$ if $i \neq j)$.

If $\bar{\xi}_i$ are the covariant components of a motion of V_n into itself, $\bar{\xi}_i$ must satisfy

* Presented to the American Mathematical Society, September 6th, 1928.

† National Research Fellow.

‡ Cf. L. P. Eisenhart, *Riemannian Geometry*, Chap. VI. It is of course understood that the summation convention of a repeated index is used throughout.

$$\bar{\xi}_{i;j} + \bar{\xi}_{j;i} = 0 \tag{1.4}$$

where a subscript preceded by a semi-colon indicates covariant differentiation with respect to $\bar{g}_{ij}$. Now

$$\begin{aligned}\bar{\xi}_{i;j} &= (\partial\bar{\xi}_i/\partial x_j) - \bar{\xi}_h \overline{\left\{ {h \atop ij} \right\}} \\ &= (\partial\bar{\xi}_i/\partial x_j) - \bar{\xi}_h \left\{ {h \atop ij} \right\} - \bar{\xi}_j\phi_{,j} - \bar{\xi}_i\phi_{,j} \\ &= \bar{\xi}_{i,j} - \bar{\xi}_j\phi_{,i} - \bar{\xi}_i\phi_{,j}.\end{aligned}$$

Therefore equations (1.4) may be written as

$$(\bar{\xi}_{i,j} - 2\bar{\xi}_i\phi_{,j}) + (\bar{\xi}_{j,i} - 2\bar{\xi}_j\phi_{,i}) = 0$$

or

$$(\bar{\xi}_i e^{-2\phi})_{,j} + (\bar{\xi}_j e^{-2\phi})_{,i} = 0. \tag{1.5}$$

If we assume that the space V_n admits an r-parameter group of motions whose covariant components are $\xi_i{}^{(\alpha)}$, $\alpha = 1, \cdots, r$, an index in a parenthesis indicating the vector, then equation (1.5) admit just r linearly independent solutions which may be taken to be

$$\bar{\xi}_i{}^{(\alpha)} = a_\beta{}^\alpha \xi_i{}^{(\beta)} e^{2\phi} = a_\beta{}^\alpha \xi_i{}^{(\beta)} (\bar{g}/g)^{1(n+1)}. \tag{1.6}$$

This proves

THEOREM I. *Every Riemannian space in geodesic correspondence with one which admits a given r-parameter continuous group of motions also admits an r-parameter continuous group of motions; the components of this group are linear combinations (with constant coefficients) of the components of the given group, multiplied by a scalar.*

As an immediate corollary of this theorem we have, since the only spaces admitting an $n(n+1)/2$-parameter group are spaces of constant curvature,

The only spaces that are in geodesic correspondence with a space of constant curvature are spaces of constant curvature,

a theorem due to Beltrami.*

2. We next suppose that V_n admits an r-parameter $(r < n)$ group of motions and that $\bar{V}_n$ is conformal to V_n; i.g. $\bar{g}_{ij} = e^{2\sigma} g_{ij}$.

If $\xi_{(\alpha)}{}^i$, $\alpha = 1, \cdots, r$, are the contravariant components of the motions in V_n they must satisfy

* Cf. L. P. Eisenhart, *loc. cit.*, p. 134.

$$g_{ih}(\partial \xi_{(a)}{}^h / \partial x_j) + g_{hj}(\partial \xi_a{}^h / \partial x^i) + \xi_{(a)}{}^h (\partial g_{ij} / \partial x^h) = 0, \tag{2.1}$$

which are Killing's equations in terms of contravariant components, and if $\xi_{(a)}{}^i$ are to be components of a motion in $\bar{V}_n$ we must have

$$\xi_{(a)}{}^h (\partial \sigma / \partial x^h) = 0, \tag{2.2}$$

as follows from (2.1). We designate the operator $\xi_{(a)}{}^h (\partial / \partial x^h)$ by X_a; then since X_a are the generators of a group we have

$$(X_a X_\beta - X_\beta X_a)\sigma = c_{a\beta}{}^\gamma X_\gamma \sigma$$

$c_{a\beta}{}^\gamma$ being the structural constants of the group. Hence the system of differential equations (2.2) is complete and admits $n - r$ functionally independent integrals.

Hence we have

THEOREM II. *If a Riemannian space admits an r-parameter $(r < n)$ group of motions there exist $n - r$ functionally independent conformal spaces admitting the same group of motions.*

PRINCETON UNIVERSITY.

Determination of All Normal Division Algebras in Thirty-six Units of Type R_2.*

By A. ADRIAN ALBERT.†

1. *Introduction.* An element x of a normal division algebra A in n^2 units over a non-modular field F is said to have grade r if its minimum equation has degree r. If there exist k distinct elements $x, \theta_1(x), \cdots, \theta_{k-1}(x)$ which are polynominals in x and satisfy its minimum equation then x is said to have type R_k. Algebra A in n^2 units has type R_k if it contains an element x of grade n and type R_k.

In all of the papers ‡ on the determination of normal division algebras the algebras were first shown to be of type R_2 and then proved to be known algebras. For algebras A in n^2 units, n a prime, the assumption that A has type R_2 implies, by an almost trivial argument, that A has type R_n and is a cyclic algebra. But when n is composite the problem of showing A to be known is difficult even under the assumption that A has type R_2. We consider here the case $n = 6$ and show that all normal division algebras in thirty-six units of type R_2 are of type R_3 and are known in the broad sense recently defined.§ A necessary and sufficient condition that the known algebras be of type R_6 and hence of the kind constructed by L. E. Dickson is also given.

2. *Results presupposed.*¶ We shall assume the following known results for normal division algebras in thirty-six units over F.

THEOREM 1. *Let x be an element of A. Any transform txt^{-1} of x by an element t of A is a root of the minimum equation of x and conversely, any root in A of the minimum equation of x is a transform of x.*

THEOREM 2. *Let $B = CD$ where C and D are polynomials in w with*

* Presented to the Society, December 27, 1928.

† National Research Fellow.

‡ Cf. [illegible] Vol. 31 (1929), pp. [illegible]-260.

§ [illegible] Vol. 36 (1929), pp. 322-339.

¶ [illegible]

coefficients in A. *Let* $\omega - y$ *be a right divisor of* B *and let* $D = Q(\omega - y) + R$ *where* $R \neq 0$ *is free of* ω. *Then* $\omega - RyR^{-1}$ *is a right divisor of* C.

Definition. x is said to have type S_k if x has type R_k and not type R_{k+1}.

THEOREM 3. *If* x *has type* S_k *and grade* r *then* k *is a divisor of* r.

THEOREM 4. *The grade of any element* x *of* A *is* 1, 2, 3 *or* 6.

THEOREM 5. *Let* x *be an element of* A *having grade six. Then the only elements of* A *commutative with* x *are polynomials in* x *with coefficients in* F.

THEOREM 6. *Let* x *be in* A *and have grade* r *and minimum equation* $\phi(\omega) = 0$. *Then there exist transforms* $x_1 = x, x_2, x_3, \cdots, x_r$ *of* x *such that*

$$\phi(\omega) \equiv (\omega - x_r)(\omega - x_{r-1}) \cdots (\omega - x_1). \tag{1}$$

Also the factors in $\phi(\omega)$ *may be permuted cyclically.*

THEOREM 7. *Let* A *contain either a cyclic sub-algebra of order nine over* F *or a generalized quaternion algebra over* F. *Then* A *has type* R_6.

THEOREM 8. *Let* A *contain* x *of grade six and type* S_2. *Then* $F(x)$ *contains an element* y *of grade six and with minimum equation*

$$\phi(\omega) \equiv \omega^6 + \alpha\omega^4 + \beta\omega^2 + \gamma = 0 \quad (\alpha, \beta, \gamma \text{ in } F). \tag{2}$$

3. *Algebras containing an element of grade* 3 *commutative with one of its transforms.* Let q_1 have grade 3 and minimum equation

$$\phi(\omega) \equiv \omega^3 + \alpha\omega^2 + \beta\omega + \gamma = 0 \quad (\alpha, \beta, \gamma \text{ in } F). \tag{3}$$

Let $q_2 = tq_1t^{-1} \neq q_1$ be a transform of q_1 and such that $q_2q_1 = q_1q_2$. Let first q_2 be a polynomial in q_1 with coefficients in F. Then (3) is a cyclic equation with

$$\theta(q_1) = tq_1t^{-1}, \ \theta^2(q_1) = t^2q_1t^{-2}, \ \theta^3(q_1) = t^3q_1t^{-3} = q_1$$

and the elements

$$q_1^i t^j (i, j = 0, 1, 2) \tag{4}$$

are linearly independent with respect to $F(t^3)$ and form a cyclic algebra T of order 9 over $F(t^3)$. The order of T with respect to F is $9a$ where a is the order of $F(t^3)$ and is a divisor of thirty-six, the order of A. If t^3 is in F then T is a cyclic algebra of order nine over F and A has type R_6 by theorem 7. Since t has at most grade six the order of T is at most eighteen, and if $F(t^3)$ has order greater than unity it is a quadratic field. Let t^3 be not in F and let r generate $F(t^3)$ such that $r^2 = \xi$ in F. Then $x = rq_1$ has grade six

since $x^2 = \xi q_1^3$ is in the cubic field $F(q_1)$ and is not in F, while x is not in $F(q_1)$ since r is not in $F(q_1)$, a cubic field thus having no quadratic subfield $F(r)$. The element x has type R_6 since if the minimum equation of ξq_1^2 is $\psi(\omega) = 0$, then the minimum equation of x is $\psi(\omega^2) = 0$ with roots rq_1, rq_2, $r(-\alpha - q_1 - q_2)$, $-rq_1$, $-rq_2$, $-r(-\alpha - q_1 - q_2)$ all of which are polynomials in x. In this case A has type R_6.

Next let q_2 be not a polynomial in q_1. Let $q_3 = -\alpha - q_1$. Consider the field $F(q_1, q_2)$. Its order is greater than three since q_2 is not in $F(q_1)$ a cubic sub-field of $F(q_1, q_2)$, so that $F(q_1, q_2)$ has order six and contains an element x of grade six generating it. Since x is not in $F(q_1)$, the elements

$$(5) \qquad 1,\ q_1,\ q_1^2,\ x,\ q_1 x,\ q_1^2 x$$

are linearly independent and form a basis of $F(x)$, so that every element of $F(x)$ is expressible in the form $a + bx$ where a and b are in $F(q_1)$. In particular $x^2 = c + dx = b - a^2 - 2ax$ for $\alpha = -2a$, $c = b - a^2$ and thus $(x + a)^2 = b$. Let $s = x + a$. We have $s^2 = b$ in $F(q_1)$ while $s = x + a$ is not in $F(q_1)$.

Let $b = \epsilon$ in F. Let $\mu_1 = sq_1$. Then $\mu_1^2 = \epsilon q_1^2$ an element of grade 3 in $F(q_1)$ while μ_1 is not in $F(q_1)$ and thus $F(\mu_1)$ has order greater than three and μ_1 has grade six. Also if the minimum equation of μ_1^2 is $\psi(\omega) = 0$ then the minimum equation of μ_1 is $\psi(\omega^2) = 0$. The field $F(\mu_1)$ is equal to $F(x)$ since s and q_1 are both in $F(x)$ while the orders of $F(x)$ and $F(\mu_1)$ are the same. Thus $\mu_2 = sq_2$ and $-\mu_2$ are both polynomials in μ_1. But $\mu_2^2 = \epsilon q_2^2$ is a transform of μ_1^2 and is a root of $\psi(\omega) = 0$ so that μ_2 and $-\mu_2$ are roots of the minimum equation of μ_1. Thus $\psi(\omega^2) = 0$ has the roots μ_1, μ_2, $-\mu_1$, $-\mu_2$. These elements are distinct since if $\mu_2 = \pm \mu_1$ then $q_2 = \pm q$, contrary to the hypothesis that q_2 is not a polynomial in q_1. Hence μ_1 has grade 6 and type R_4 and, by theorem 3 has type R_6 so that A has type R_6.

Consider next b not in F. Then b has grade three and since $s^2 = b$ and s is not in $F(b) = F(q_1)$, the grade of s is greater than three and is six. Also if the minimum equation of b is $\psi(\omega) = 0$ then the minimum equation of s is $\psi(\omega^2) = 0$ and $-s$ is a root of it. The elements

$$(6) \qquad 1,\ q_1,\ q_1^2,\ q_2,\ q_1 q_2,\ q_1^2 q_2$$

are linearly independent and form a basis of $F(q_1, q_2) = F(s)$ since q_2 is not in the cubic field $F(q_1)$. Thus there exist elements g and h in $F(q_1)$ such that $s = g + hq_2$. Let $s_2 = g(q_2) + h(q_2)$ [illegible]. We may write [illegible] in the form

$$(7) \qquad \phi(\omega) \equiv (\omega - q_3)(\omega - q_2)(\omega - q_1).$$

Comparing coefficients in (3) and (7) we obtain

$$(8) \qquad q_3 = -(q_1 + q_2 + \alpha), \quad q_3q_2 + q_3q_1 + q_2q_1 = \beta.$$

Substitute the value of q_3 in the second equation of (8). We obtain the relation

$$(9) \qquad q_2^2 = -(q_1 + \alpha)q_2 - (\beta + q_1^2 + \alpha q_1).$$

But $s^2 = g^2 + 2ghq_2 + h^2q_2^2 = b$ in $F(q_1)$. Thus $-h^2(q_1 + \alpha) + 2gh = 0$. Since s is not in $F(q_1)$, $h \neq 0$ and $2g = h(q_1 + \alpha)$, so that $2s = h(q_1 + \alpha + 2q_2)$, and

$$\begin{aligned} 4s^2 &= h^2[4q_2^2 + 4(q_1 + \alpha)q_2 + q_1^2 + 2\alpha q_1 + \alpha^2] \\ &= h^2[-4(\beta + q_1^2 + \alpha q_1) + q_1^2 + 2\alpha q_1 + \alpha^2] = 4\sigma(q_1). \end{aligned}$$

By formally replacing q_1 by q_2 and q_2 by q_3 in the above we obtain $s_2^2 = \sigma(q_2)$. But $\sigma(q_2)$ is a transform of $\sigma(q_1) = b$ and is a root of $\psi(\omega) = 0$. Thus s_2 is a root of $\psi(\omega^2) = 0$. If $s_2 = \pm s$, then $s_2^2 = s_1^2 = \sigma(q_1) = \sigma(q_2)$. Since b is not in F and thus generates $F(q_2)$ this implies that q_2 is a polynomial in $\sigma(q_2) = \sigma(q_1)$ and hence in q_1, a contradiction of our hypothesis. Thus $\psi(\omega^2) = 0$ has the distinct roots s_2, $-s_2$, s, $-s$ and s has type R_4. Thus s has type R_6 and A has type R_6. This proves the following theorem.

THEOREM 9. *Let A be a normal division algebra in thirty-six units containing an element q of grade three which is commutative with one of its transforms not equal to itself. Then A has type R_6.*

4. *Algebras containing an element q_1 commutative with no one of its transforms.* Let q_1 have grade 3 and minimum equation

$$(10) \qquad \phi(\omega) \equiv \omega^3 + \alpha\omega^2 + \beta\omega + \gamma = 0 \qquad (\alpha, \beta, \gamma \text{ in } F).$$

We may write

$$(11) \qquad \phi(\omega) \equiv (\omega - q_3)(\omega - q_2)(\omega - q_1) = 0,$$

where q_3, q_2, q_1 are transforms of q, and $q_3q_1 \neq q_1q_3$, $q_2q_1 \neq q_1q_2$. Comparing coefficients in (10) and (11) we have

$$(12) \qquad q_3q_2 + q_3q_1 + q_2q_1 = \beta.$$

The factors in (11) may be permuted cyclically and we may show that $x = q_3q_2 - q_2q_3 = q_1q_3 - q_3q_1 = q_2q_1 - q_1q_2 \neq 0$ by means of (12) and

$$(13) \qquad q_2q_1 + q_2q_3 + q_1q_3 = q_1q_3 + q_1q_2 + q_3q_2 = p.$$

Applying theorem 2 with $y = q_2$, $B = CD \equiv \phi(\omega)$, $C = \omega - q_3$, $D = (\omega - q_2)(\omega - q_1)$, $R = q_2 q_1 - q_1 q_2 \neq 0$, we show that $x q_2 x^{-1} = q_3$ and similarly by cylic permutation of the factors of (11), $x q_3 x^{-1} = q_1$, $x q_1 x^{-1} = q_2$, so that $x^3 q_1 x^{-3} = q_1$ and $x^3 = d$ is commutative with each of q_1, q_2, q_3. Let S be the algebra composed of all polynomials in q_1 and q_2 with coefficients in F. S contains x and thus q_3. The quantity d is commutative with each element of S. Let K be the field composed of all elements of S commutative with every element of S. Then d is in K and S is a normal division algebra over K.

LEMMA 1. *Let $\Gamma(\omega) = 0$ be the minimum equation of x with respect to K. Then, if $x_2 \equiv (\alpha + 3q_2) x (\alpha + 3q_2)^{-1}$, and we have the decomposition*

$$\Gamma(\omega) \equiv Q(\omega)(\omega - x_1), \tag{14}$$

$\omega - x_2$ is a right divisor of $Q(\omega)$.

Proof. $\omega - (q_1 - q_2) x (q_1 - q_2)^{-1}$ is a right divisor of $\Gamma(\omega) = 0$ since $q_1 q_2 \neq q_2 q_1$ and thus $q_1 - q_2 \neq 0$. It follows that if

$$\begin{aligned} R &= (q_1 - q_2) x (q_1 - q_2)^{-1} - x \\ &= [(q_1 - q_2) x - x (q_1 - q_2)] (q_1 - q_2)^{-1} \\ &= (q_1 - q_2 - q_2 + q_3) x (q_1 - q_2)^{-1} \neq 0, \end{aligned}$$

then $\omega - R(q_1 - q_2) x (q_1 - q_2)^{-1} R^{-1}$ is a right divisor of $Q(\omega)$. But $q_1 + q_3 = -\alpha - q_2$, so that $R = -(\alpha + 3q_2) x (q_1 - q_2)^{-1} \neq 0$ since q_2 is not in F, and $R(q_1 - q_2) x (q_1 - q_2)^{-1} R^{-1} = (\alpha + 3q_2) x (\alpha + 3q_2)^{-1}$, proving the lemma.

LEMMA 2. *x has not grade 2 with respect to K.*

Proof. Let x be a root of

$$\psi(\omega) \equiv \omega^2 + \xi\omega + \eta \qquad (\xi, \eta \text{ in } K). \tag{15}$$

Then

$$\psi(\omega) = (\omega + \xi + x)(\omega - x).$$

Applying lemma 1 we have

$(\alpha+3q_2)^2$ and in $(\alpha+3q_3)^2$. Thus $q_2q_3=q_3q_2$ a contradiction since $x=q_3q_2-q_2q_3\neq 0$, and the lemma is proved.

Let x_2 be defined as in lemma 1. Then

$$\psi(\omega)\equiv\omega^3-d\equiv(\omega-x_3)(\omega-x_2)(\omega-x_1), \tag{16}$$

where $x_1=x$, x_3 is some transform of x. We shall first consider the case $x_1x_2\neq x_2x_1$ and consequently $y=x_2x_1-x_1x_2\neq 0$.

By comparing coefficients in (16) we may show that $y=x_3x_2-x_2x_3$ $=x_1x_3-x_3x_1$. We also have $yx_1y^{-1}=x_2$, $yx_2y^{-1}=x_3$, $yx_3y^{-1}=y^3x_1y^{-3}=x_1$ and y^3 is commutative with x_1, x_2, x_3. Let $z_1=x_1y$, $z_2=x_1z_1x_1^{-1}=x_1^2yx_1^{-1}$ and $y^3=\nu$. Let T be the algebra composed of all polynomials in x_1 and z_1. Then T contains $y=x_1^{-1}z$, $x_2=yx_1y^{-1}$, $x_3=yx_2y^{-1}$. The element $d=x_1^3$ is commutative with x_1 and x_2 and consequently with $y=x_2x_1-x_1x_2$ and $z_1=x_1y$. Hence d and ν are both commutative with every element of T. Let H be the field of all elements of T commutative with every element of T. The field H contains d and ν and T may be considered as a normal division algebra over H. By (16), $x_3^3=d=x_3x_2x_1$ and $x_3^2-x_2x_1=0$. But

$$z_1z_2-z_2z_1=x_1(yx_1^2y)x_1^{-1}-x_1^2y^2=x_1(yx_1^2y-x_1y^2x_1)x_1^{-1}$$

and

$$\begin{aligned}x_3^2-x_2x_1&=y^2x_1^2y^{-2}-yx_1y^{-1}x_1\\&=yx_1^2(1/\nu)y-yx_1y^2(1/\nu)x_1=y(yx_1^2y-x_1y^2x_1)(1/\nu)=0\end{aligned}$$

so that $yx_1^2y-x_1y^2x_1=0$ and $z_1z_2=z_2z_1$ since $y\neq 0$ and $1/\nu\neq 0$ is commutative with x_1 and y. Let $z_3=x_1^2z_1x_1^{-2}=x_1^3yx_1^{-2}=yx_1$. Then $z_1z_3-z_3z_1$ $=x_1y^2x_1-yx_1^2y=0$ and $z_2z_3-z_3z_2=x_1(z_1z_2-z_2z_1)x_1^{-1}-0$. But, since $z_1=x_1y$, $z_1^2=x_1x_2y^2$,

$$z_1^3=x_1x_2y^2x_1y=x_1x_2x_3\nu=\nu(x_2x_1-y)x_3=\nu(x_2x_1x_3)-\nu x_1y=\nu(d-z_1),$$

z is a root of

$$\omega^3+\nu\omega-\nu d=0. \tag{17}$$

LEMMA 3. *The element z_1 has grade 3 with respect to H.*

Proof. The grade of z_1 with respect to H is <3 since z_1 satisfies (17). Now z_1 has not grade 1 with respect to H since z_1 is not commutative with x_1, an element of T. Let z_1 satisfy

$$P(\omega)\equiv\omega^2+\lambda\omega+\mu=0 \qquad (\lambda,\ \mu \text{ in } H). \tag{18}$$

Then $P(\omega)\equiv(\omega+\lambda+z_1)(\omega-z_1)$. But $\omega-z_2$ is a right divisor of and is commutative with $(\omega-z_1)$ so that $(\lambda+z_1)=-z_2=-x_1z_1x_1^{-1}$,

$$x_1z_2x_1^{-1}=z_3=-x_1(\lambda+z_1)x_1^{-1}=-\lambda x_1x_1^{-1}+\lambda+z=z_1,$$
$$z_3=yx_1=z_1=x_1y,$$

a contradiction. This proves the lemma.

Hence the minimum equation of z_1 with respect to H is

$$\psi(\omega)\equiv\omega^3+\nu\omega-\nu d=0. \tag{19}$$

Consider the field $F(z_1, z_2, z_3)$. If its order is not six then it is three by lemma 3, and z_2 and z_3 are transforms of z_1 which are commutative with z_1, an element of grade 3. Hence, by the work of section 3, A has type R_6. If $F(z_1, z_2, z_3)$ has grade 6 then there exists an element t of grade six in $F(z_1, z_2, z_3)$ generating the field, and t has type R_3 since $x_1tx_1^{-1}$, $x_1^2tx_1^{-2}$, t are distinct and are in $F(t)=F(z_1, z_2, z_3)$.

It remains to consider the case $x_2x_1=x_1x_2$. If x_1 has grade three then, since x_2 is a transform of x_1 commutative with x_1, A has type R_6 by the work of the preceding section. If x_1 has not grade three then, since its grade with respect to H is three, its grade with respect to F is greater than three and is six. Thus x_2 is a polynomial in x and $x_3x_1=x_1x_3$ so that x_3 is a polynomial in x_1 and x_3, x_2, x_1 are all roots of the minimum equation of x_1 and are distinct by (16) and lemma 1. Hence x_1 has grade 6 and type R_3 and we have obtained the theorem

THEOREM 10. *Let A be a normal division algebra in thirty-six units containing an element of grade three. Then A has type R_3.*

5. *Algebras of type R_2.* Let A have type R_2. Then A contains an element x of grade six and type R_2 so that x has type R_3 or type S_2. If x has type S_2 then, by theorem 8, $F(x)$ and consequently A contains an element y of grade six and minimum equation $\phi(\omega^2)=0$, so that y^2 has grade 3 and, by theorem 10, A has type in R_3. This proves

THEOREM 11. *Every normal division algebra in thirty-six units of type R_2 has type R_3.*

In the author's paper on "Algebras in $4p^2$ Units, p an odd Prime" (*loc. cit.*), it was shown that all normal division algebras of type R_3 were known, that is contained an element i whose minimum equation

$$\phi(\omega) \equiv \omega^6 + \alpha\omega^4 + \beta\omega^3 + \gamma\omega^2 + \delta\omega + \epsilon = 0 \qquad (\alpha, \beta, \gamma, \delta, \epsilon \text{ in } F) \tag{20}$$

had type R_6 or S_3. If (15) has type S_3, that is there exists a polynomial in i, $\theta(i)$, such that $i = \theta^3(i)$, $\theta(i)$, $\theta^2(i)$ are roots of (15) and no other polynomial in i is a root of 15, then A was shown to have the following structure:

A contains Σ which is a cyclic algebra of order 9 over the quadratic field $F(q)$ of all polynomials in i which are symmetric in the functions $i, \theta(i), \theta^2(i)$. The algebra Σ has a basis

$$1, i, i^2, y, iy, i^2y, y^2, iy^2, i^2y^2 \tag{21}$$

and the multiplication table given by (15) for $\omega = i$, and

$$y^\lambda f(i) = f[\theta^\lambda(i)]y^\lambda \qquad (\lambda = 0, 1, 2), \quad y^3 = \gamma \text{ in } F(q) \tag{22}$$

for every $f(i)$ of $F(i)$, and the quadratic field $F(q)$ is generated by q such that $q^2 = \xi$ in F. There exists an element k in Σ and not in $F(i)$ and an element y' in Σ such that

$$q(k) = -q(i), \quad (y')^3 = \gamma(-q), \quad (y')^r f(k) = f[\theta^r(k)](y')^r \qquad (r = 0, 1, 2), \tag{23}$$

so that if we write for any element

$$X = a(i) + b(i)y + c(i)y^2, \quad X' = a(k) + b(k)y' + c(k)y'^2 \tag{24}$$

then we have the properties

$$\alpha' = \alpha, \quad (X + Y)' = X' + Y', \quad (XY)' = X'Y' \tag{25}$$

for every X and Y of Σ and α of F. Also there exists an element s in Σ such that if we write $X^{('')} = (X')'$, $X^{(''')} = (X'')', \cdots$, then

$$s' = s, \quad X^{('')} = sXs^{-1}. \tag{26}$$

Then there exists an element z in A such that the elements

$$X_1 + X_2 z \tag{27}$$

where X_1 and X_2 range independently over all elements of Σ give all elements of A and A has the multiplication table

$$zX = X'z, \quad z^2 = s. \tag{28}$$

Finally, the only elements of A commutative with q are elements of Σ.

Conversely let i satisfy (20) and let there exist an element y such that the elements (21) and their multiplication table (22) are given and form an associative algebra Σ. Then Σ is a division algebra if and only if $\gamma \neq a[\theta^2(x)]a[\theta(x)]a(x)$ for any polynomial $a(x)$. Let elements k, y' and s of Σ be given so that (23) and (26) are satisfied. Then it is known that there exists an element z such that the elements (27) form the elements of an algebra A of type R_3 in 36 units with multiplication table (28). Also A is an associative normal algebra and is a division algebra if and only if $s \neq x'x$ for any x of Σ.

We shall now consider algebras B in thirty-six units of type R_6. Let B be such an algebra and let B contain x of type R_6.

LEMMA 4. *Let Ω be a substitution group of order six on six letters. Then Ω has a cyclic sub-group of order* 3 *and a cyclic sub-group of order* 2.

Proof. It is known * that the only groups of order and degree six are the cyclic group and one further group. This second group contains one sub-group of order three and three sub-groups of order two. The cyclic group may be written $\{I, \Theta, \Theta^2, \Theta^3, \Theta^4, \Theta^5\}$ where $\Theta^6 = I$. This contains the sub-group $\{I, \Theta^2, \Theta^4\}$ of order three and the sub-group $\{I, \Theta^3\}$ of order 2.

Hence if x has grade six and type R_6, since its minimum equation has a regular group, we may write the roots x, $\theta_1(x)$, $\theta_1{}^2(x)$, $\theta_2(x)$, $\theta_2[\theta_1(x)] = \theta_1[\theta_2(x)]$, $\theta_2[\theta_1{}^2(x)] = \theta_1{}^2[\theta_2(x)]$ such that $\theta_1{}^3(x) = \theta_2{}^2(x) = x$. Let first $\theta_2(x) + x = h \neq 0$. Let $g = x - \theta_2(x)$. If $g^2 = \xi$ in F, $h^2 = \eta$ in F, then either the field $F(g, h)$ has a basis $1, g, h, gh$ and hence has order 4 which is impossible since the grade of any element of B is 1, 2, 3, or 6, or $F(g, h) = F(g)$ of order two and h is in $F(g)$ and is unaltered by replacing $\theta_2(x)$ by x and x by $\theta_2(x)$, contrary to its form. Hence either ξ is not in F or η is not in F. Let $p = gh$. Then $-p$ is a root of the minimum equation of i since if we write $p = \theta(x)$ then if $\psi(p) = 0$ this equation is unaltered by replacing x by $\theta_2(x)$ and consequently g by $-g$. Hence the minimum equation of p has even powers only and p has grade two or six. If p has grade two then $p^2 = \zeta$ in F. But the minimum equation of g has even powers only so that either η is in F or g has grade six. But not both of η and ζ are in F since not both of ξ and η are in F and $\zeta = \zeta\eta$. Therefore one of i and g has grade six and minimum equation $\psi(\omega^2) = 0$. Also if we call i the element g or p according as g or p has grade six then $F(i) = F(x)$ and $i = \theta(x)$.

---

* See Netto, *Theory of Substitutions*, pp. 146-148.

Thus $\theta[\theta_1(x)]$, $\theta[\theta_1^2(x)]$ are roots of $\psi(\omega^2)=0$ and we obtain the following theorem

THEOREM 12. *Let x have grade 6 and type R_6. Then $F(x)$ contains an element i and a polynomial $\pi(i)$ such that the minimum equation of i is*

$$(29)\quad \psi(\omega)\equiv\omega^6+\lambda\omega^4+\mu\omega^2+\nu$$
$$\equiv[\omega+\pi^2(i)][\omega-\pi^2(i)][\omega+\pi(i)][\omega-\pi(i)][\omega+i][\omega-i]=0,$$
$$\lambda,\ \mu,\ \nu \text{ in } F.$$

It is obvious that i^2 has grade three and that $[\pi^2(i)]^2$, $[\pi(i)]^2$, i^2 are roots of its minimum equation commutative with it so that, by theorem 9,

THEOREM 13. *Let A be a normal division algebra in thirty-six units. Then A has type R_6 if and only if it contains an element of grade three commutative with one of its transforms not equal to itself.*

The algebras of type R_6 are those of the kind constructed by L. E. Dickson in his Transactions (1926) paper. These algebras are known algebras of type R_3 with the properties (22)-(26) where now, if we take i as the element of type R_6 and (29) as its minimum equation, we have $k=-i$. We have also found a necessary and sufficient condition that known algebras of type R_3 be of type R_6, in theorem 13.

PRINCETON, N. J.

On the r-th Divisors of a Number.

By D. H. LEHMER.

Introduction. Associated with every positive integer n is its set of distinct divisors (including 1 and n). Let us designate this set by δ_1. We have also a set of divisors associated with each number in the set δ_1. In this way the set δ_1 gives rise to a new set δ_2 of numbers (not all distinct) which are in fact the divisors of the divisors of n itself. If we take the divisors of the numbers in the set δ_2 we obtain a new set δ_3 and so on. The elements of the set δ_r we call the rth divisors of n. If $n > 1$ it is clear that the number of elements in the set δ_r increases with r. The order in which the elements in each set are arranged is of no consequence.

With the number n are also associated its proper divisors, those divisors of n which are less than n. This set of numbers we call d_1. Each number (> 1) in this set has one or more proper divisors. The set of all proper divisors of the numbers in the set d_1 we call d_2. The proper divisors of the numbers in the set d_2 form a set d_3 and so on. The numbers in the set d_r we call the rth proper divisors of n. The number of rth proper divisors of n ultimately becomes zero although it may increase for the first few values of r. It is convenient to think of sets δ_0 and d_0 each consisting of a single element, namely n itself.

The following example illustrates the nature of these sets of numbers in the case $n = 24$, and will be of use later.

$\delta_1 = 1, 2, 3, 4, 6, 8, 12, 24.$

$\delta_2 = 1 \quad 1, 2 \quad 1, 3 \quad 1, 2, 4 \quad 1, 2, 3, 6 \quad 1, 2, 4, 8 \quad 1, 2, 3, 4, 6, 12$
$\quad 1, 2, 3, 4, 6, 8, 12, 24.$

$\delta_3 = 1 \quad 1 \quad 1, 2 \quad 1 \quad 1, 3 \quad 1 \quad 1, 2 \quad 1, 2, 4 \quad 1 \quad 1, 2 \quad 1, 3 \quad 1, 2, 3, 6$
$\quad 1 \quad 1, 2 \quad 1, 2, 4 \quad 1, 2, 4, 8 \quad 1 \quad 1, 2 \quad 1, 3 \quad 1, 2, 4 \quad 1, 2, 3, 6$
$\quad 1, 2, 3, 4, 6, 12 \quad 1 \quad 1, 2 \quad 1, 3 \quad 1, 2, 4 \quad 1, 2, 3, 6 \quad 1, 2, 4, 8$
$\quad 1, 2, 3, 4, 6, 12 \quad 1, 2, 3, 4, 6, 8, 12, 24.$

. . . .

$d_1 = 1, 2, 3, 4, 6, 8, 12.$

$d_2 = 1 \quad 1 \quad 1, 2 \quad 1, 2, 3 \quad 1, 2, 4 \quad 1, 2, 3, 4, 6.$

$d_3 = 1 \quad 1 \quad 1 \quad 1 \quad 1, 2 \quad 1 \quad 1 \quad 1, 2 \quad 1, 2, 3.$

$d_4 = 1 \quad 1 \quad 1 \quad 1$

For $r > 4$ the sets d_r are empty.

We first study the connection between the same arbitrary function of δ_r and of d_r which leads to certain inversion formulas. These we apply to the determination of the product and the sum of the kth powers of the rth divisors of n. Solutions of certain problems in the theory of probability can be obtained as a by-product of this investigation. If r is negative the definition of the rth divisors of n is meaningless. Nevertheless if we put negative values of r in the expressions for the above sums we obtain a sequence of functions different in nature but essentially belonging to the same class. Among these new functions we find Euler's totient function $\phi(n)$ and Merten's inversion function $\mu(n)$. In fact the whole investigation can be considered as a simultaneous extension of these two classical functions which thus appear to be closely connected.

Inversion of Numerical Integrals. The following investigation is simplified by the introduction of the concept of numerical integration. If two numerical functions $F(n)$ *and* $f(n)$ *are connected by the relation*

$$F(n) = \Sigma f(\delta_1) \tag{1}$$

where δ_1 ranges over the divisors of n, then $F(n)$ is said to be the numerical integral of $f(n)$; conversely $f(n)$ is the numerical derivative of $F(n)$. This concept is due to Bougaief * and has led to remarkable relations existing between the standard discontinuous functions of the theory of numbers. There are various inversion formulas by means of which we can express $f(n)$ in (1) in terms of $F(n)$. The most compact expression is perhaps

$$f(n) = \Sigma\mu(n/\delta_1)F(\delta_1) \tag{2}$$

which is due to Laguerre.† The function $\mu(n)$ is Merten's inversion function which is zero if n contains a square factor and $(-1)^t$ if n is a product of t distinct primes and is unity if $n=1$. Another form of (2) is due to Dedekind: ‡

$$f(n) = F(n) - \Sigma F(n/p_1) + \Sigma F(n/p_1p_2) - \Sigma F(n/p_1p_2p_3) + \cdots$$

where $n = p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_t^{\alpha_t}$. We assume throughout the article that n is expressed in this form. We also make use of the quantity

$$\omega = \alpha_1 + \alpha_2 + \cdots + \alpha_t,$$

which is called the order of n.

* *Matematicheskii Sbornik*, Vol. 5 (1870-72), pp. 1-63; *Bulletin des Sciences Mathématiques et Astronomiques*, Vol. 10, pp. 13-32.

† *Bulletin de la Société Mathématique de France*, Vol. 1 (1872-3), pp. 77-81.

‡ *Journal für Mathematik*, Vol. 54 (1857), p. 21.

Introducing our sets d_r we can give another expression for $f(n)$ as follows:

$$(4) \qquad f(n) = F(n) - \Sigma F(d_1) + \Sigma F(d_2) - \Sigma F(d_3) + \cdots (-1)^{\omega} \Sigma F(d_{\omega}).$$

To prove this formula we take for values of n the numbers in the set d_1 and n itself and sum. The result is

$$\Sigma f(d_1) + f(n) = \Sigma F(d_1) - \Sigma F(d_2) + \Sigma F(d_3) - \Sigma F(d_4) + \cdots$$
$$+ F(n) - \Sigma F(d_1) + \Sigma F(d_2) - \Sigma F(d_3) + \Sigma F(d_4) - \cdots .$$

Hence $\Sigma f(\delta_1) = F(n)$.

Formula (4) gives us new expressions for certain important functions. For example, if $\phi(n)$ is Euler's totient function we can write, in view of Gauss' theorem $n = \Sigma\phi(\delta_1)$, that

$$\phi(n) = n - \Sigma(d_1) + \Sigma(d_2) - \Sigma(d_3) + \cdots .$$

THEOREM 1. *Let $F(n)$ be an arbitrary function of n then for every integer r we have*

$$(5) \qquad (-1)^r \Sigma F(d_r) = F(n) - \binom{r}{1} \Sigma F(\delta_1) + \binom{r}{2} \Sigma F(\delta_2) - \cdots (-1)^r \Sigma F(\delta_r).$$

In fact $\Sigma F(d_1) = \Sigma F(\delta_1) - F(n)$.

To obtain $\Sigma F(d_2)$ we sum over the proper divisors of δ_1 and n, namely over the set $\delta_2 - \delta_1$ and the set $\delta_1 - n$. That is,

$$\Sigma F(d_2) = \Sigma F(\delta_2) - 2\Sigma F(\delta_1) + F(n).$$

Similarly

$$\Sigma F(d_3) = \Sigma F(\delta_3) - 3\Sigma F(\delta_2) + 3\Sigma F(\delta_1) - F(n).$$

From the relation

$$(6) \qquad \binom{m-1}{l-1} + \binom{m-1}{l} = \binom{m}{l}$$

we have by a simple induction

$$\Sigma F(d_r) = \Sigma F(\delta_r) - \binom{r}{1} \Sigma F(\delta_{r-1}) + \binom{r}{2} \Sigma F(\delta_{r-2}) - \cdots + (-1)^r F(n),$$

which is the theorem. In the same way it can be shown that

$$(7) \qquad \Sigma F(\delta_r) = \Sigma F(d_r) + \binom{r}{1} \Sigma F(d_{r-1}) + \binom{r}{2} \Sigma F(d_{r-2}) + \cdots + F(n).$$

If we sum equations of the type (5) for $r = 0, 1, 2, \cdots, \omega$ we obtain in view of (4)

$$(8) \qquad f(n) = \binom{\omega+1}{1} F(n) - \binom{\omega+1}{2} \Sigma F(\delta_1)$$
$$+ \binom{\omega+1}{3} \Sigma(\delta_2) - \cdots + (-1)^{\omega} \Sigma F(\delta_{\omega}).$$

which gives us a new inversion formula. As a matter of fact, if the index r in (5) exceeds ω, then $\Sigma F(d_r) = 0$. So that in (8) we may take in place of ω any larger number and the result will be the same. As an example of (8) we may put $f(n) = n$ and get another expression for Euler's totient function in the form

$$(9) \qquad \phi(n) = \tbinom{r}{1}n - \tbinom{r}{2}\Sigma\delta_1 + \tbinom{r}{3}\Sigma\delta_2 - \cdots + (-1)^{r-1}\Sigma\delta_{r-1},$$

where $r \geq \omega + 1$.

For instance, if $n = 6$ we have

$\delta_1 = 1, 2, 3, 6.$
$\delta_2 = 1, \quad 1, 2, \quad 1, 3, \quad 1, 2, 3, 6.$
$\delta_3 = 1, \quad 1, \quad 1, 2, \quad 1, \quad 1, 3, \quad 1, \quad 1, 2, \quad 1, 3, \quad 1, 2, 3, 6.$
$\delta_4 = 1, \quad 1, \quad 1, \quad 1, 2, \quad 1, \quad 1, \quad 1, 3, \quad 1, \quad 1, \quad 1, 2, \quad 1, \quad 1, 3, \quad 1, \quad 1, 2, \quad 1, 3, \quad 1, 2, 3, 6.$
$\Sigma\delta_1 = 12, \quad \Sigma\delta_2 = 20, \quad \Sigma\delta_3 = 30, \quad \Sigma\delta_4 = 42.$

Applying (9) we have for $r = 3, 4, 5$,

$$\phi(6) = 3.6 - 3.12 + 20 = 2,$$
$$\phi(6) = 4.6 - 6.12 + 4.20 - 30 = 2,$$
$$\phi(6) = 5.6 - 10.12 + 10.20 - 5.30 + 42 = 2.$$

If F and f are related not as in (1) but as follows

$$(1') \qquad F(n) = \Pi f(\delta_1),$$

or what is the same

$$\log F(n) = \Sigma \log f(\delta_1),$$

the above inversion formulas (4) and (8) go over into

$$(4') \qquad f(n) = \frac{F(n) \cdot \Pi F(d_2) \cdot \Pi F(d_4) \cdots}{\Pi F(d_1) \cdot \Pi F(d_3) \cdots}$$

$$(8') \qquad f(n) = \frac{[F(n)]^{\binom{\omega+1}{1}} \cdot [\Pi F(\delta_2)]^{\binom{\omega+1}{3}} \cdots}{[\Pi F(\delta_1)]^{\binom{\omega+1}{2}} \cdot [\Pi F(\delta_3)]^{\binom{\omega+1}{4}} \cdots}$$

Successive Numerical Integrals. Let us consider a sequence of functions

$$(10) \qquad \cdots f_{-3}, \quad f_{-2}, \quad f_{-1}, \quad f_0, \quad f_1, \quad f_2, \quad f_3, \cdots$$

each of which is the numerical integral of its predecessor. That is

$$(11) \qquad f_r(n) = \Sigma f_{r-1}(\delta_1).$$

If we define one of the functions, say $f_0(n)$, for all values of n, then every function $f_r(n)$ is completely determined.

These functions f_r have certain properties in common. For instance

$$(12) \qquad f_r(1) = f_0(1) \qquad \text{for every } r.$$

The following multiplicative property is important for our purposes:

$$(13) \qquad f(m) \cdot f(n) = f(mn),$$

where m and n are relatively prime integers.

THEOREM 2. *If any function f_r of the set* (10) *possesses the multiplicative property* (13), *so does every function of the set.*

The proof is by induction. We shall first show that f_{r+1} has the property. From (11) we have

$$(14) \qquad f_{r+1}(n) \cdot f_{r+1}(m) = \sum_{\delta_1, \delta_1'} f_r(\delta_1) \cdot f_r(\delta_1'),$$

where δ_1' ranges over the divisors of m. Since m and n are prime to each other and since (13) holds for f_r, the right side of (14) becomes $\Sigma f_r(\delta_i\delta_1')$. As δ_1 and δ_1' run over their respective values their product ranges over the divisors of mn without repetition. Hence

$$f_{r+1}(n) \cdot f_{r+1}(m) = f_{r+1}(mn).$$

To prove that (13) holds for f_{r-1} we make use of (2) and write

$$f_{r-1}(n) \cdot f_{r-1}(m) = \sum_{\delta_1, \delta_1'} \mu(n/\delta_1) f_r(\delta_1) \cdot \mu(n/\delta_1') f_r(\delta_1')$$
$$= \Sigma\, \mu(mn/\delta_1\delta_1') f_r(\delta_1\delta_1'),$$

since $\mu(n)$ has the multiplicative property.* Thus

$$f_{r-1}(n) \cdot f_{r-1}(m) = f_{r-1}(mn).$$

Hence the theorem is proved.

The expression (11) can be generalised at once to give the relation between f_r and f_{r+s}.

$$(15) \qquad \begin{aligned} f_{r+1}(n) &= \Sigma f_r(\delta_1), \\ f_{r+2}(n) &= \Sigma f_{r+1}(\delta_1) = \Sigma f_r(\delta_2), \\ &\cdots\cdots\cdots\cdots \\ f_{r+s}(n) &= \Sigma f_r(\delta_s) = \Sigma f_s(\delta_r). \end{aligned}$$

Comparing (15) and (7) we have

$$(16) \qquad f_{r+s}(n) = f_r(n) + \tbinom{s}{1} \Sigma f_r(d_1) + \tbinom{s}{2} \Sigma f_r(d_2) + \cdots + \Sigma f_r(d_s).$$

This expression for the integral of order s of f_r is to be preferred to (15) for computational purposes. After the finite number of sums

* This simple fact follows at once from the definition of $\mu(n)$.

$$\Sigma f_r(d_i) \qquad (i=0, 1, 2, 3, \cdots, \omega)$$

have been calculated, the integral of any order may be obtained as the linear combination of these with the proper binomial coefficients.

Example. Let $f_r(n) = P_0(n)$, the number of primes less than or equal to n. The problem of getting an explicit formula even for the first integral is a very difficult one. Nevertheless we can easily calculate the numerical value of the integral of any order for a given n. For instance for $n = 24$ referring to the sets d_r already given in the introduction, we have the following values for the sums:

$$P_0(n) = 9, \quad \Sigma P_0(d_1) = 17, \quad \Sigma P_0(d_2) = 15, \quad \Sigma P_0(d_3) = 5, \quad \Sigma P_0(d_4) = 0.$$

Hence

$$P_s(24) = 9 + 17\binom{s}{1} + 15\binom{s}{2} + 5\binom{s}{3}.$$

We can express $f_r(n)$ in (15) in terms of $f_{r+s}(n)$ by making successive applications of (2) with the result

$$f_r(n) = \Sigma\, \mu(n/\delta_s) f_{r+s}(\delta_s), \tag{17}$$

which gives us a way of calculating isolated derivatives of a given function f_{r+s}. Successive applications of the formula (4) give an expression similar to (16), namely

$$\begin{aligned} f_r(n) = f_{r+s}(n) - \binom{s}{1} \Sigma f_{r+s}(d_1) + \binom{s+1}{2} \Sigma f_{r+s}(d_2) - \cdots \\ + (-1)^\omega \binom{\omega+s-1}{\omega} \Sigma f_{r+s}(d_\omega). \end{aligned} \tag{18}$$

Putting $n = 24$ and $s = 5$ we calculate $P_5(\delta_1)$ from (16) with the result

$$P_5(24) = 294, \quad P_5(12) = 85, \quad P_5(8) = 29, \quad P_5(6) = 18,$$
$$P_5(4) = 7, \quad P_5(3) = 2, \quad P_5(2) = 1.$$

Using these values in (18) we have

$$P_0(24) = 294 - \binom{5}{1} \cdot 145 + \binom{6}{2} \cdot 40 - \binom{7}{3} \cdot 5 = 9,$$

which is, in fact, the number of primes less than 24.

In view of relation (15) we can transform (5) to read

$$\begin{aligned} (-1)^s \Sigma f_r(d_s) &= f_r(n) - \binom{s}{1} f_{r+1}(n) + \binom{s}{2} f_{r+2}(n) - \cdots (-1)^s f_{r+s}(n) \\ &= \sum_{i=0}^{s} (-1)^i \binom{s}{i} f_{r+i}(n). \end{aligned} \tag{19}$$

This last sum is zero for all values of $s > \omega$ regardless of what sequence of functions f_r we select.

In the same way the inversion formula (8) may be written

$$(20)\qquad f_r(n) = \binom{h}{1} f_{r+1}(n) + \binom{h}{2} f_{r+2}(n) + \binom{h}{3} f_{r+3}(n) - \cdots + (-1)^{h-1} f_{r+h}(n),$$

where $h > \omega$.

The Function $\sigma_{k,r}(n)$. We shall consider as a special case of (10) the sequence $\sigma_{k,r}(n)$ determined by putting $f_0(n) = n^k$. From (15) it follows that

$$\sigma_{k,r}(n) = \Sigma\, \sigma_{k,0}(\delta_r) = \Sigma\, (\delta_r)^k,$$

which shows that $\sigma_{k,r}(n)$ is the sum of the k-th powers of the r-th divisors of n.

THEOREM 3. *If p is a prime, then*

$$\sigma_{k,r}(p^a) = p^{ka} + \binom{r}{1} p^{k(a-1)} + \binom{r+1}{2} p^{k(a-2)} + \cdots = \sum_{j=0}^{a} \binom{j+r-1}{j} p^{k(a-j)}.$$

The theorem is evidently true for $r = 0$. The proof of the theorem consists in showing that if it is true for the index r, it is also true for both the indices $r - 1$ and $r + 1$, the first part of the proof being required for the discussion of negative indices. First we may write

$$(21)\qquad \begin{aligned} \sigma_{k,r}(p^a) &= \sigma_{k,r-1}(p^a) + \sum_{j=0}^{a-1} \sigma_{k,r-1}(p^j), \\ \sigma_{k,r-1}(p^a) &= \sigma_{k,r}(p^a) - \sigma_{k,r}(p^{a-1}). \end{aligned}$$

Supposing the theorem to be true for the index r we have

$$\sigma_{k,r-1}(p^a) = p^{ka} + \binom{r}{1} p^{k(a-1)} + \binom{r+1}{2} p^{k(a-2)} + \binom{r+2}{3} p^{k(a-3)} + \cdots - p^{k(a-1)} - \binom{r}{1} p^{k(a-2)} - \binom{r+1}{2} p^{k(a-3)} - \cdots.$$

In view of (6)

$$\sigma_{k,r-1}(p^a) = p^{ka} + \binom{r-1}{1} p^{k(a-1)} + \binom{r}{2} p^{k(a-2)} + \binom{r+1}{3} p^{k(a-3)} + \cdots.$$

Hence the theorem holds true for the index $r - 1$.

To prove that the theorem is true for $r + 1$ we use an induction on the exponent α. Writing (21) in the form

(22) [illegible]

[illegible]

Hence the theorem is true for $\sigma_{k,r+1}(p)$. To complete the proof we must show that if the theorem holds for $\sigma_{k,r+1}(p^{a-1})$ and $\sigma_{k,r}(p^a)$ it is true for $\sigma_{k,r+1}(p^a)$. With these hypotheses equation (22) becomes

$$\begin{aligned}\sigma_{k,r+1}(p^a) &= p^{ka} + \binom{r}{1}p^{k(a-1)} + \binom{r+1}{2}p^{k(a-2)} + \cdots \\ &\quad + p^{k(a-1)} + \binom{r+1}{1}p^{k(a-2)} + \cdots \\ &= p^{ka} + \binom{r+1}{1}p^{k(a-1)} + \binom{r+2}{2}p^{k(a-2)} + \cdots .\end{aligned}$$

This establishes the theorem for the index $r+1$ and hence for all indices.

It is obvious that $\sigma_{k,0}(n)$ has the multiplicative property (13). By Theorem 2 all σ's have this property. This leads us at once to the general expression for $\sigma_{k,r}(n)$ namely

$$(23) \qquad \sigma_{k,r}(n) = \prod_{i=1}^{t} \sum_{j=0}^{a_i} \binom{j+r-1}{j} p_i^{k(a_i-j)}.$$

The function $\mu_r(n)$. We call $\mu_r(n)$ the special case of the function $\sigma_{k,r}(n)$ for which $k=0$. From the last equation,

$$(24) \qquad \mu_r(n) = \prod_{i=1}^{t} \sum_{j=0}^{a_i} \binom{j+r-1}{j} = \prod_{i=1}^{t} \binom{r+a_i}{r}.$$

The function $\mu_r(n)$ gives the number of r-th divisors of n, that is the number of elements in the set δ_r. For $r=1$, $\mu_1(n) = \prod_{i=1}^{t}(\alpha_i+1)$, which the well known formula for the number of first divisors of n. The function $\mu_r(n)$ does not depend on the actual prime factors of n but upon their number and multiplicity. In case n is a product of t distinct primes $\mu_r(n) = (r+1)^t$.

Example. Let $n = p_1^3 p_2^2 p_3$, then $\mu_r(n) = (r+1)^3(r+2)^2(r+3)/12$, which gives for the first five values of r:

$$\begin{array}{lll}\mu_0(n) = 1, & \mu_1(n) = 24, & \mu_2(n) = 180, \\ \mu_3(n) = 800, & \mu_4(n) = 2625, & \mu_5(n) = 7056.\end{array}$$

THEOREM 4. *The number of r-th proper divisors of n is given by*

$$(25) \quad M_r(n) = \mu_r(n) - \binom{r}{1}\mu_{r-1}(n) + \binom{r}{2}\mu_{r-2}(n) - \cdots + (-1)^r \mu_0 \cdot (n).$$

This follows at once from equation (19). If n is a product of t distinct primes, (25) becomes

$$M_r(n) = (r+1)^t - \binom{r}{1}r^t + \binom{r}{2}(r-1)^t - \binom{r}{3}(r-2)^t + \cdots + (-1)^r.$$

If $r > t$ this expression is zero since the sets d_r are empty. The vanishing

of this expression for $r > t$ has been noted in connection with binomial series.*

If we introduce negative values of the index, the function $\mu_r(n)$ becomes †

$$(26) \quad \mu_{-r}(n) = \prod_{i=1}^{t} \binom{a_i - r}{a_i} = \prod_{i=1}^{t} (-1)^{a_i} \binom{r-1}{a_i} = (-1)^{\omega} \prod_{i=1}^{t} \binom{r-1}{a_i}.$$

The function $\mu_{-1}(n)$ is zero if $n > 1$ and unity if $n = 1$. The function $\mu_{-2}(n)$ is zero if n contains a square factor, is $(-1)^t$ if n is a product of t distinct primes and is unity if $n = 1$. In short $\mu_{-2}(n)$ is Merten's inversion function. This proves that if a function $f(n)$ is such that $\Sigma f(\delta_1)$ is zero for $n > 1$ and is unity for $n = 1$, then $f = \mu$.‡ The function $\mu_{-3}(n)$ vanishes if n is divisible by a cube. Otherwise its value is $(-2)^{\lambda}$ where λ is the number of prime factors of n which appear to the first power only. In general $\mu_{-r}(n)$ is zero if n is divisible by a perfect r-th power.

Example. If we take again $n = p_1^3 p_2^2 p_3$, we have

$\mu_{-1} = \mu_{-2} = \mu_{-3} = 0, \quad \mu_{-4} = 9, \quad \mu_{-5} = 96, \quad \mu_{-6} = 500, \quad \mu_{-7} = 1800.$

The function $\phi_r(n)$. We define the function

$$(27) \qquad \phi_r(n) \equiv \sigma_{1,r}(n) = \prod_{i=1}^{t} \sum_{j=0}^{a_i} \binom{j+r-1}{r-1} p_i^{a_i - j}$$

which gives the sum of the r-th divisors of n. If $r = 1$ we have

$$\phi_1(n) = \prod_{i=1}^{t} (p_i^{a_i} + p_i^{a_i - 1} + \cdots + 1) = \prod_{i=1}^{t} (p_i^{a_i+1} - 1)/(p_i - 1),$$

which is well known. For $r = 2$ we have

$$\phi_2(n) = \prod_{i=1}^{t} [p_i^{a_i+2} - p_i(\alpha_i + 2) + \alpha_i + 1]/(p_i - 1)^2,$$

which gives the sum of the numbers in the set δ_2. The function

$$\phi_3(n) = \prod_{i=1}^{t} \left[p_i^{a_i+3} - \binom{a_i+3}{2} p_i^2 + (\alpha_i + 1)(\alpha_i + 3) p_i - \binom{a_i+2}{2}\right] / (p_i - 1)^3,$$

gives the sum of the third divisors of n, and so on.

For negative indices, (27) becomes §

$$\phi_{-r} = \prod \sum (-1)^{j} [illegible] \; ([illegible]).$$

[illegible]

If $r = 1$ we have

$$\phi_{-1} = \prod_{i=1}^{t} (p^{a_i} - p^{a_i - 1}),$$

which is Euler's totient function.* This converse of Gauss' theorem has been given by Lucas.† The functions ϕ_{-r} for $r > 1$ have certain similarities to Euler's $\phi(n)$. For instance

$$(28) \qquad \begin{aligned} \phi_{-r}(p^{a}) &= p^{a-r}(p-1)^{r} \qquad \alpha \geqq r \\ \phi_{-r}(p) &= p - r. \end{aligned}$$

THEOREM 5. *If n contains the prime factor p to the first but no higher power, then $\phi_{-p}(n) = 0$.*

This follows at once from (28).

The following is a list of values of $\phi_r(n)$ for $n = 42$ and $n = 56$

$r =$	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3
$\phi_r(42) =$	-168	-84	-30	0	12	12	6	0	0	12	42	96	180	300
$\phi_r(56) =$	186	80	24	0	-6	-4	0	4	10	24	56	120	234	420

The function $\pi_r(n)$. In conclusion we shall develop a formula for the product $\pi_r(n)$ of the r-th divisors of n. We shall use the following lemma

LEMMA. The number of times that a certain divisor Δ of n appears in the set δ_r is $\mu_{r-1}(n/\Delta)$.

Evidently the lemma is true for $r = 1$ since δ_1 consists of distinct elements and $\mu_0 = 1$. Let us assume that the lemma is true for the set δ_{r-1}. If we form the set δ_r from the divisors of the set δ_{r-1} the element Δ can arise only from numbers of the form $\Delta\delta'$ where δ' are the divisors of (n/Δ) and from each of these it can arise but once. By the hypothesis of our induction the number of times each divisor of the form $\Delta\delta'$ appears in δ_{r-1} is $\mu_{r-2}(n/\Delta\delta')$. The total number of Δ's appearing in δ_r is therefore

$$\Sigma\mu_{r-2}(n/\Delta\delta') = \Sigma\mu_{r-2}(\delta') = \mu_{r-1}(n/\Delta)$$

since δ' ranges over the divisors of (n/Δ). Thus the lemma is true for all positive values of r.

Let $f_s(n)$ be any function of an arbitrary sequence. Then from the lemma it follows that

$$(29) \qquad f_{r+s}(n) = \Sigma f_s(\delta_r) = \Sigma\mu_{r-1}(n/\delta_1) f_s(\delta_1).$$

---

* In general $\sigma_{k,-1}(n) = n^k \prod_{i=1}^{t} (1 - P_i^{-k})$. This function has been the subject of a number of investigations. See D. N. Lehmer, *American Journal of Mathematics*, Vol. 22 (1900), pp. 293-335.

† *Théorie des Nombres*, p. 401.

This gives another short method of calculating the actual values of isolated integrals and derivatives. In this method it is unnecessary to determine the numbers in the sets δ_k for $k > 1$. Thus to find $P_5(n)$ in Example 1 we put $f_s(n) = P_0(n)$ and $r = 5$. The results are as follows

$$\begin{array}{rcccccccccc}
\delta_1 = & 1, & 2, & 3, & 4, & 6, & 8, & 12, & 24, & \\
P_0(\delta_1) = & 0, & 1, & 2, & 2, & 3, & 4, & 5, & 9, & \\
\mu_4(n/\delta_1) = & 175, & 75, & 35, & 25, & 15, & 5, & 5, & 1, & \\
P_5(n) = & 0 & +\,75 & +\,70 & +\,50 & +\,45 & +\,20 & +\,25 & +\,9 & = 294,
\end{array}$$

which agrees with our previous result.

In (29) r and s are interchangable and since we chose s arbitrarily r is similarly at our disposal that is (29) will hold for $r \leq 0$, although the expression $\Sigma f_s(d_r)$ fails to have any meaning in this case. If we put $r = -1$ we have

$$f_{s-1}(n) = \Sigma \mu_{-2}(n/\delta_1) f_s(\delta_1).$$

Since μ_{-2} is Merten's function this expression is simply Leguerre's inversion formula (2). From this standpoint μ_{-r} may be thought of as an inversion function of order $r-1$. Putting $f_s(n) = n$ in (29) we have the following compact expression for $\phi_r(n)$

$$\phi_r(n) = \Sigma \delta_1 \mu_{r-1}(n/\delta_1).$$

If we wish to find the number of times a divisor Δ appears in the set d_r we use equation (5) with $F(n) = 0$ for n different from Δ and $F(\Delta) = 1$. This gives the result

$$\mu_{r-1}(n/\Delta) - \binom{r}{1}\mu_{r-2}(n/\Delta) + \binom{r}{2}\mu_{r-3}(n/\Delta) - \cdots + (-1)^r \mu_{-1}(n/\Delta).$$

THEOREM 6. *The product $\pi_r(n)$ of the r-th divisors of n is $n^{\mu_r(n)/(r+1)}$.* Let us select a certain prime factor p of n and write $n = p^a m$. We shall determine the power to which p appears in the desired product. Let us first consider the numbers of the set δ_r which contain p to the first but not to any higher power. Such numbers are of the form pm' where m' are the divisors of m. By our lemma the number of such elements is

$$\mu_r(m)\sum_{j=1}^{a} j\mu_{r-1}(p^{a-j}) = \mu_r(m)[\mu_r(p^{a-1}) + \mu_r(p^{a-2}) + \cdots + 1]$$
$$= \mu_r(m)\,\mu_{r+1}(p^{a-1}).$$

Since
$$\mu_{r+1}(p^{a-1}) = \mu_r(p^a)\cdot\alpha/(r+1)$$

the power to which p appears in the final product is

$$[\alpha/(r+1)]\mu_r(m)\,\mu_r(p^a) = [\alpha/(r+1)]\mu_r(n).$$

This result being true for all primes dividing n we have

$$\pi_r(n) = \prod_{i=1}^{t} p_i^{\alpha_i\mu_r(n)/(r+1)} = n^{\mu_r(n)/(r+1)},$$

which is the theorem.

For $r = 1$ we have $\pi_1(n) = n^{\mu_1(n)/2}$. This well known result follows at once by associating with every divisor $\delta_1 \leqq n^{1/2}$ its co-divisor. This method fails however for $r > 1$.

The product of the numbers in the set d_r may be readily obtained from inversion formula (5) putting $\Sigma F(\delta_k) = [\mu_k(n)/(k+1)]\log n$.

Corollary. The geometric mean of the numbers in the set δ_r *is* $n^{1/(r+1)}$. This follows immediately from the theorem since $\mu_r(n)$ gives the number of elements in the set δ_r. It is interesting to notice that the mean can be calculated without reference to the prime factors of n and two numbers such as 8191 and 8192 have practically the same mean although their compositions are widely different.

Although the foregoing discussion was carried out in number-theoretic terminology it is of much wider application. We might consider a collection of ω things of which α_1 are of one sort, α_2 of a second sort and so on. It is well known that the number of selections that can be made from such a collection is the same as the number of first divisors of $n = p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_t^{\alpha_t}$. The function $\mu_r(n)$ has immediate application to a similar problem in which r successive selections are made. If we count as inadmissible the selection in which all things are rejected, the r-th proper divisors of n must be considered. If we assign different weights to the different sorts of things the function $\phi_r(n)$ becomes essential to the discussion. In a future paper we hope to apply the theory of r-th divisors to certain problems in combinatory analysis.

On Harmonics Applicable to Surfaces of Revolution.

By D. M. Wrinch.

Introduction.

1. A large class of problems in mathematical physics may be described in the following manner. A function V is required, satisfying the equation of Laplace, which is evanescent on S, the sphere at infinity, has no singularities in the region between S and a certain closed surface s, and on this surface s takes a certain specific form. This general class includes problems of electrical distributions on conductors, both when the conductor is freely charged and when its electrification is due to an external field of electric force. It further covers problems of hydrodynamics, when, for example, there is streaming past an obstacle, or when a body has uniform translational or rotational motion in an infinite liquid.

2. Problems of this type, though worked out with some degree of completeness for two dimensional surfaces,* have been solved for but a few three dimensional surfaces. Those associated with the sphere can be solved in terms of the spherical harmonics

$$r^{-k-1}\, P_k^m(\cos\theta)\cdot[\cos m\omega,\ \sin m\omega]$$

where r, θ, ω are the usual spherical polar coördinates measured from the centre of the sphere as origin. In addition to the case of the prolate and oblate spheroids, these problems have also been fully worked out for inverted oblate spheroids, the centre of inversion being on the axis of figure and they may also, evidently, be discussed for prolate spheroids, similarly inverted.† These extensions of the technique for dealing with spheroids proved possible owing to the fact that if harmonics are available for any surface they are also available for the inverse surface. Ellipsoidal harmonics, further, make the treatment of problems associated with ellipsoids theoretically possible,

---

* See a series of papers by D. M. Wrinch: *Philosophical Magazine*, Vol. 48 (1924), pp. 692-703 and pp. 1089-1104; also Vol. 49 (1925), pp. 240-250; *Journal of the Royal Aeronautical Society*, Vol. 30 (1926), pp. 129-141.

† See papers by D. M. Wrinch in the *Philosophical Magazine*, Vol. 53 (1927), pp. 865-883; and Vol. 50 (1925), pp. 1049-1058.

though the practical difficulties are so great that very little progress has, in fact, been made.

3. It appears, however, that there has been very little methodological treatment of these problems, whereby they may be attacked even when the surface is of a general type. It seemed worth while, therefore, to put forward a method of constructing harmonics which will allow treatment of problems in which there is symmetry about an axis, when associated with a surface of revolution whose equation is

$$(1)\quad \begin{aligned} z &= z(u) = \alpha \cos u + \alpha_1 \cos u + \alpha_2 \cos 2u + \cdots + \alpha_n \cos nu + \cdots \\ \rho &= \rho(u) = \alpha \sin u - \alpha_1 \sin u - \alpha_2 \sin 2u - \cdots - \alpha_n \sin nu - \cdots \end{aligned}$$

(the axis of symmetry being the axis of figure), provided that the coefficients α, α_n satisfy certain general conditions. Thus, in the first place, we shall assume that the transformation

$$(2)\qquad W = W(w) = \alpha e^{iw} + \alpha_1 e^{-iw} + \alpha_2 e^{-2iw} + \cdots + \alpha_n e^{-niw} + \cdots$$

where $w = u + iv$ is such that dW/dz is never zero and never infinite for values of v between zero and $-\infty$. The zero v level evidently represents the surface s, and the v level given by $v = -\infty$ the sphere at infinity: the assumption then ensures that dW/dz has no zeros and no infinities in the region between s and S, the sphere at infinity. A further condition relating to the orders of magnitude of the coefficients will also be introduced in due course.

4. The harmonics to be constructed are interesting in that they are all simply series of Legendre Functions $P_k(\cos u)$ multiplied by polynomials in e^v. This fact renders the method of some practical importance, since the Legendre functions are already adequately tabulated and easy to handle. It also has the consequence that much of the theory of Legendre functions already worked out will undoubtedly prove useful and illuminating in the treatment of surfaces which possess a more general equation than the sphere and spheroid previously treated.

5. But the chief interest of the method is of a different kind. In the case of the sphere, given by the zero v level in the transformation $W = \alpha e^{iw}$, problems involving solutions symmetric about the z axis are solved in terms of the harmonics

$$P_k(\cos u)\ e^{(k+1)v}$$

If we go one step on the road towards generality and treat the spheroid by means of the transformation

$$W = \alpha e^{iw} + \alpha_1 e^{-iw}, \tag{3}$$

(in which $\alpha_1^2 < \alpha^2$) there is a sense in which the harmonics for the sphere are 'corrected' into harmonics for the spheroid, of type

$$P_k(\cos u)\ F_k(e^v),$$

although the customary treatment by means of spheroidal harmonics has obscured the fact.

Now there is also a sense in which harmonics available for any surface of the type given in (1) are 'corrected' forms of spheroidal harmonics, but this procedure has the disadvantage that the standard forms for the harmonics would then involve the Q_k and q_k Legendre functions, which do not lend themselves readily to simple computation. We prefer therefore to take the spherical harmonics as the starting point and to build up harmonics for more general surfaces from them.

The advantage of this point of view is that it enables us to look at the problems as a whole, whatever type of surface of revolution is under discussion, provided only that it falls under the equation for some set of values or other of the α_n parameters which they are free to take. The way to obtain harmonics for any special surface is theoretically plain and in any particular case presents no mathematical difficulties of any sort. Indeed, it is not difficult to elaborate a technique whereby the form of harmonic suitable for a surface s can be 'corrected' as we pass from the spherical surface treated by $W = \alpha e^{iw}$ through intermediate stage to the greater generality covered by the transformation (2) in its various cases. The interesting point is that there is an inductive procedure always available for these problems whereby a more general form for the surface s does not require a wholly new treatment *ab initio*, but merely the correction of the harmonics in certain clearly defined ways. It is not necessary to wait for a kind of mathematical mutation before a new type of surface can be treated.

involved are arithmetically complicated. Since it is no part of our plan in this paper to work out pages of arithmetical calculations, having worked out a few particular cases and having explained the method and its scope, we shall leave on one side the laborious determination of the coefficients involved in other still more general cases. It may, perhaps, turn out to be the case that new results relating to Legendre functions will be required before any complete account of them can be given.

6. It will be noticed that there is a considerable advantage from the point of view of applications in the fact that a step by step method of building harmonics is available. In practice it is of the first importance that results should be expressed in such a form that an approximation to any required degree is immediately deducible. Indeed it is sometimes the case that elegant solutions of problems in pure mathematics give only cold comfort to those who wish to use them to obtain further information about the external world. We therefore think it important to draw attention to the fact that the harmonics developed in this paper are of a type which allows the solution of problems of physics to be pursued to just that degree of approximation which the particular circumstances of each problem may require.

I. *The appropriate form for Laplace's Equation.*

1.1. When

$$x,\, y = \rho(\cos\omega,\, \sin\omega)$$
$$z + i\rho = W(w) = W(u + iv)$$

we take the line elements normal to the u, v, ω levels in the form

$$dn_u = du/h_u,\ dn_v = dv/h_v,\ dn_\omega = d\omega/h_\omega$$

and in the usual manner

$$1/h^2_{u,v,\omega} = (\partial x/\partial u, v, \omega)^2 + (\partial y/\partial u, v, \omega)^2 + (\partial z/\partial u, v, \omega)^2$$

so that

$$1/h^2_\omega = \rho^2$$
$$1/h^2_u = 1/h^2_v = |\, dw/dz\, |^2 = 1/h^2$$

In consequence, Laplace's equation takes the form

$$\nabla^2 V \equiv h^2 \left[\frac{1}{\rho}\ \frac{d}{du}\, \rho\, \frac{d}{du} + \frac{1}{\rho}\ \frac{d}{dv}\, \rho\, \frac{d}{dv} + \frac{1}{\rho^2 h^2}\ \frac{d^2}{d\omega^2} \right] V = 0,$$

and solutions symmetrical about the axis of z, being independent of ω, therefore have the determining equation

$$\left[\frac{d^2}{du^2} + \frac{1}{\rho}\,\frac{d\rho}{du}\,\frac{d}{du} + \frac{d^2}{dv^2} + \frac{1}{\rho}\,\frac{d\rho}{dv}\,\frac{d}{dv}\right] V = 0.$$

Now with

$$Z + i\rho - W(u + iv) = \alpha e^{iw} + \alpha_1 e^{-iw} + \alpha_2 e^{-2iw} + \cdots + \alpha_n e^{-niw} + \cdots$$

as in (2), the equation takes the form

$$(4) \qquad \phi(u, v)\left[(d^2/du^2) + (d^2/dv^2)\right]V + \psi(u, v)\,(dV/du) + \chi(u, v)\,(dV/dv) = 0$$

where

$$\begin{aligned} \phi &\equiv (\alpha e^{-v} - \alpha_1 e^{v}) \sin u - \alpha_2 e^{2v} \sin 2u \cdots \\ \psi &\equiv (\alpha e^{-v} - \alpha_1 e^{v}) \cos u - 2\alpha_2 e^{2v} \cos 2u \cdots \\ -\chi &\equiv (\alpha e^{-v} + \alpha_1 e^{v}) \sin u + 2\alpha_2 e^{2v} \sin 2u \cdots \end{aligned}$$

which may be written in terms of operators D, D_n

$$(5) \qquad (\alpha D - \alpha_1 D_1 - \alpha_2 D_2 \cdots)\, V = 0$$

when the operators are defined by the equations

$$D = \frac{1}{\sin u}\,\frac{d}{du}\,\sin u\,\frac{d}{du} + \frac{d^2}{dv^2} - \frac{d}{dv}$$

$$D_n = e^{(n+1)v}\left\{\frac{1}{\sin u}\,\frac{d}{du}\,\sin nu\,\frac{d}{du} + \frac{\sin nu}{\sin u}\left(\frac{d^2}{dv^2} + n\,\frac{d}{dv}\right)\right\}$$

We may then regard the equation (5) as the standard form of Laplace's equation for a surface of revolution given by (1).

It will be convenient to use, also, the co-ordinates ζ, μ defined by

$$(6) \qquad e^{v} = \zeta, \qquad \cos u = \mu.$$

With this definition $\zeta = 1$ is the surface s and $\zeta = 0$ the sphere at infinity [illegible]

[illegible]

[illegible]

and
$$\theta_\zeta \equiv \zeta(d/d\zeta)$$

(8) $$D_n \equiv \zeta^{n+1}\{(d/d\mu)\chi_n(\mu)(1-\mu^2)(d/d\mu) + \chi_n(\mu)\theta_\zeta(\theta_\zeta + n)\}$$
with

(9) $$\chi_n(\mu) \equiv \sin n\,(\cos^{-1}\mu)/\sin(\cos^{-1}\mu).$$

Then $\chi_n(\mu)$ is evidently a series of terms in μ^{n-1}, μ^{n-3}, $\mu^{n-5}\cdots$ and in particular

(10) $$\chi_1 = 1,\ \chi_2 = 2\mu,\ \chi_3 = 4\mu^2 - 1.$$

II. *The Standard Form for General Harmonics.*

2. 1. The fundamental case to deal with is when $\alpha_1, \alpha_2 \cdots \alpha_n \cdots$ are all zero. s is then the sphere $r = \alpha$, being in fact the zero v level in the transformation

$$z + i\rho = \alpha e^{iu-v}$$

or the unit ζ level in the transformation

$$z = \alpha \cos u/\zeta, \quad \rho = \alpha \sin u/\zeta.$$

The equation for V becomes simply

$$DV = 0;$$

and since $P_k(\mu)$ satisfies the equation

(11) $$[(d/d\mu)(1-\mu^2)(d/d\mu) + k(k+1)]P_k(\mu) = 0,$$

and ζ^{k+1} the equation

(12) $$[\theta_\zeta(\theta_\zeta - 1) - (k+1)k]\zeta^{k+1} = 0;$$

a harmonic evanescent on S, the sphere at infinity, is therefore available in the form

$$V = P_k(\mu)\zeta^{k+1}$$

recognisable in the usual form

$$V = P_k(\mu), \quad (\alpha/r)^{k+1}$$

since $r = \alpha/\zeta$. We can then write down a solution of Laplace's equation for the sphere which on the spherical surface s is equal to any function of z and ρ, or (as for example in the case of the problem of the freely charged conductor) is constant on s.

2.2. Now it will be convenient to systematize the surface s to be treated by assuming that if we write

$$\alpha_n/\alpha = \gamma_n \lambda^n, \qquad (\lambda^2 < 1)$$

then the coefficients

$$\gamma_1, \gamma_2 \cdots \gamma_n \cdots$$

are all of the same order of magnitude. It will also be possible to deal with the more general case when there exist integers $i(n)$ such that if we write

$$\alpha_n/\alpha = \gamma_n \lambda^{i(n)} \qquad (\lambda^2 < 1)$$

then the γ coefficients are all of the same order of magnitude. Thus we might have

$$(\alpha_1, \alpha_2)/\alpha = (\gamma_1, \gamma_2)\lambda; \quad (\alpha_3, \alpha_4)/\alpha = (\gamma_3, \gamma_4)\lambda^2 \cdots$$

and the method now under discussion could easily be made applicable. However, it is not feasible to treat generally the cases when the functions $i(n)$ are required, though each separate case can readily be treated. It must always be the case that the basic assumption as to dW/dz is still operative, and this will evidently limit the γ coefficients in certain well defined ways.

Suppose now, under the assumption that the γ coefficients are all of the same order of magnitude, we take

$$(13) \qquad V = f_0 + \lambda f_1 + \cdots + \lambda^n f_n + \cdots$$

and solve the equations

$$(14) \qquad \begin{aligned} Df_0 &= 0 \\ Df_1 &= \gamma_1 D_1 f_0 \\ Df_2 &= \gamma_1 D_1 f_1 + \gamma_2 D_2 f_0 \\ Df_3 &= \gamma_1 D_1 f_2 + \gamma_2 D_2 f_1 + \gamma_3 D_3 f_0 \end{aligned}$$

and in general

$$Df_{m+1} = \gamma_1 D_1 f_m + \gamma_2 D_2 f_{m+1} + \cdots + \gamma_{m+1} D_{m+1} f_0.$$

Then we can develop a harmonic to any degree of approximation in the form

$$v = f_0 + \lambda f_1 \cdots + \lambda^p f_p + O(\lambda^{p+1})$$

and the exact form of any harmonic will be

$$v = f_0 + \lambda f_1 \cdots + \lambda^p f_p + \cdots.$$

To find any harmonic we require therefore to be able to deal with the equations (14).

2. 3. Since D is the usual operator for spherical polar co-ordinates, these equations have the characteristic that the complementary functions are always simply

$$\Sigma a_p P_k(\mu)\zeta^{p+1}.$$

We may also explicitly notice that

$$DP_m(\mu)\zeta^\sigma = (\sigma - m - 1)(\sigma + m)P_m(\mu)\zeta^\sigma \tag{15}$$

in the usual way, for this fact has the important consequence that if in any equation

$$Dv = F(\mu, \zeta)$$

F consists of the sum of such terms as $b_{m,\sigma}P_m(\mu)\zeta^\sigma$ (in which $\sigma - m - 1 \neq 0 \neq \sigma + m$), then a particular integral is available in the form

$$\Sigma c_{m,\sigma}P_m(\mu)\zeta^\sigma$$

where

$$c_{m,\sigma} = b_{m,\sigma}/(\sigma - m - 1)(\sigma + m)$$

and a general solution involving an enumerable infinity of arbitrary constants is given by

$$\Sigma a_p P_p(\mu)\zeta^{p+1} + \Sigma c_{m,\sigma}P_m(\mu)\zeta^\sigma. \tag{16}$$

There is thus no theoretical difficulty whatever in solving any equation when the right side F consists of P functions multiplied by polynomials in ζ.

Now we easily find that

$$D_1 P_m(\mu)\zeta^\sigma = (\sigma - m)(\sigma + m + 1)\zeta^{\sigma+2}P_m(\mu) \tag{17}$$

and we shall shortly show that, in general,

$$D_n P_m(\mu)\zeta^\sigma = \zeta^{\sigma+n+1}[(\sigma, m)^n_{-n+1}P_{m-n+1} + (\sigma, m)^n_{-n+3}P_{m-n+3} + \cdots + (\sigma, m)^n_{n-3}P_{m+n-3} + (\sigma, m)^n_{n-1}P_{m+n-1}]. \tag{18}$$

Then it follows by an inductive argument, that since

$$f_0 = \Sigma g_k \zeta^{k+1}P_k(\mu)$$

the right side in all the equations (14) is of the requisite form for the simple solutions given in (16).

Thus a harmonic in the general case now under discussion takes the standard form

$$v = \Sigma \Pi^{m}(\zeta) P_p(\mu)$$

where by $\Pi^{m}(x)$ we mean a polynomial function of x of degree m.

It will be a simple matter in the sequel to investigate the degrees of the polynomials $\Pi(\zeta)$ and to show how in any particular case their actual coefficients can be evaluated.

III. *A Study of the D_n Operators.*

3.1. The operator D_n is given by

$$\begin{aligned} D_n &= \zeta^{n+1}\{(d/d\mu)\chi_n(\mu)(1-\mu^2)(d/d\mu) + \chi_n(\mu)\theta_\zeta(\theta_\zeta + n)\} \\ &= \zeta^{n+1}\{\chi_n(\mu)[(d/d\mu)(1-\mu^2)(d/d\mu) + \theta_\zeta(\theta_\zeta + n)] \\ &\quad + \chi_n'(\mu)(1-\mu^2)(d/d\mu)\} \end{aligned}$$

where

$$\chi_n(\mu) = \sin nu/\sin u = s_{n-1}\mu^{n-1} + s_{n-3}\mu^{n-3} + \cdots.$$

Now since

$$(d/d\mu)(1-\mu^2)(d/d\mu)P_p = -p(p+1)P_p$$

and

$$(2p+1)(1-\mu^2)(d/d\mu)P_p = p(p+1)(P_{p-1} - P_{p+1})$$

therefore

$$\begin{aligned} D_n\zeta^\sigma P_m(\mu) = \zeta^{\sigma+n+1}\ &\{\chi_n(\mu)[-m(m+1) + \sigma(\sigma+n)]P_m \\ &+ \chi_n'(\mu)\cdot[m(m+1)/(2m+1)](P_{m-1} - P_{m+1})\}. \end{aligned}$$

It is also an important characteristic of P_m functions that

$$(2p+1)\mu P_p(\mu) = pP_{p-1}(\mu) + (p+1)P_{p+1}(\mu),$$

for it then follows that $\mu^2 P_p$ gives terms in P_{p-2}, P_p, P_{p+2} and in general that $\mu^q P_p$ gives terms in P_{p-q}, $P_{p-q+2} \cdots P_{p+q-2}$, P_{p+q}. Evidently therefore $\chi_n(\mu)P_m$ may be expressed in terms of

$$P_{m-(n-1)},\ P_{m-(n-3)} \cdots P_{m+(n-3)},\ P_{m+(n-1)};$$

and, in a similar fashion, $\chi_n'(\mu)P_{m\pm1}$ may be expressed in terms of

[illegible]

[illegible]

[illegible]

It will prove a severe exercise in Legendre functions to work out the coefficients $(\sigma, m)_r{}^n$ in general. However any required coefficients can be obtained and we append the value of members of this class of coefficient which will allow a number of applications of the method to be carried out. Thus in addition to the expression for D_1 already given in (17), we give

$$D_2\zeta^\sigma P_m(\mu) = \zeta^{\sigma+3}[(\sigma, m)_{-1}{}^2 P_{m-1} + (\sigma, m)_{+1}{}^2 P_{m+1}]$$

with

$$(2m+1)(\sigma, m)_{-1}{}^2 = 2m(\sigma - m + 1)(\sigma + m + 1)$$
$$(2m+1)(\sigma, m)_{+1}{}^2 = 2(m+1)(\sigma - m)(\sigma + m + 2)$$

and

$$D_3\zeta^\sigma P_m(\mu) = \zeta^{\sigma+4}[(\sigma, m)_{-2}{}^3 P_{m-2} + (\sigma, m)_0{}^3 P_m + (\sigma, m)_2{}^3 P_{m+2}]$$

with

$$(2m-1)(2m+3)(\sigma, m)_0{}^3 = \sigma(\sigma+3)(4m^2+4m-1) - m(m+1)(4m^2+4m-9)$$

and

$$(2m+1)(2m-1)(\sigma, m)_{-2}{}^3 = 4m(m-1)(\sigma - m + 2)(\sigma + m + 1)$$
$$(2m+1)(2m+3)(\sigma, m)_2{}^3 = 4(m+1)(m+2)(\sigma - m)(\sigma+m+3).$$

IV. *The Construction of General Harmonics.*

4.1. It is now possible to work out harmonics satisfying Laplace's equation in its general form

$$(D - \gamma_1\lambda D_1 - \gamma_2\lambda^2 D_2 \cdots - \gamma_n\lambda^n D_n \cdots)v = 0.$$

The equation for f_0 allows a general solution

$$\Sigma\, a_k P_k(\mu)\zeta^{k+1}.$$

But let us investigate the harmonic—which it will be convenient to call $\Phi_k(\mu, \zeta)$—in which

$$f_0 = P_k(\mu)\zeta^{k+1}.$$

In solving the equations (14) for $f_1, f_2 \cdots f_n \cdots$ which are all of the form

$$Df_n = \Sigma\, b_{m,\sigma} P_m(\mu)\zeta^\sigma,$$

we shall uniformly select the particular integral

$$f_n = \Sigma\, c_{m,\sigma} P_m(\mu)\zeta^\sigma$$

where

$$c_{m,\sigma} = b_{m,\sigma}/(\sigma - m - 1)(\sigma + m)$$

and add no term of the complementary function given in (16). In this way we shall build up a harmonic $\Phi_k(\mu, \zeta)$ in the simplest possible manner. Nevertheless no generality is lost, for by introducing at any stage—say in the solution of the equation for f_n—any term of the complementary function, as for example

$$f_n = \Sigma\, c_{m,\sigma} P_m(\mu)\zeta^\sigma + a_p P_p(\mu)\zeta^{p+1}$$

we shall ultimately arrive at the solution

$$\nu = \Phi_k(\mu, \zeta) + a_p \lambda^n \Phi_p(\mu, \zeta)$$

and so obtain the sum of multiples of two standard harmonics, instead of one alone.

4. 2. Adopting then the solution

$$_kf_0 = P_k \zeta^{k+1}$$

we have

$$\nu = \Phi_k(\mu, \zeta) = P_k \zeta^{k+1} + O(\lambda)$$

and since the equation for f_1 is

$$Df_1 = \gamma_1 D_1 f_0 = 2\gamma_1 (k+1) P_k \zeta^{k+3}$$

we therefore take

$$_kf_1 = [(k+1)/(2h+3)]\, \gamma_1 P_k \zeta^{k+3}.$$

Now the equation for f_2 is

$$Df_2 = \gamma_2 D_2 f_0 + \gamma_1 D_1 f_1$$

and as

$$(2k+1) D_2 P_k \zeta^{k+1} = 4\zeta^{k+4}[4k(k+1)P_{k-1} + (k+1)(2h+3)P_{k+1}]$$
$$D_1 P_k \zeta^{k+3} = 6(k+2)\zeta^{k+5} P_k$$

we adopt the particular integral

$$_kf_2 = \zeta^{k+4}[c_k P_k \zeta + c_{k-1} P_{k-1} + c_{k+1} P_{k+1}]$$

with

[illegible]

[illegible]

[illegible]

[illegible]

$$Df_3 = \gamma_3 D_3 f_0 + \gamma_2 D_2 f_1 + \gamma_1 D_1 f_2$$

and the facts that

$$D_3 P_k \zeta^{k+1} = \zeta^{k+5}[(k+1,k)_{-2}{}^3 P_{k-2} + (k+1,k)_0{}^3 P_k + (k+1,k)_2{}^3 P_{k+2}]$$

$$D_2 P_k \zeta^{k+3} = 2\zeta^{k+6}[4k(2h+4)P_{k-1} + 3(k+1)(2h+5)P_{k+1}]/(2k+1)$$

$$D_1(P_{k-1}, P_{k+1})\zeta^{k+4} = \zeta^{k+6}[5(2h+4)P_{k-1}, 3(2k+6)P_{k+1}]$$

we deduce the particular integral

$$\begin{aligned} {}_kf_3 = (k+1)\zeta^{k+5}\Bigg[\Bigg\{ & \frac{4k^2+6k-1}{(2k-1)(2h+3)(2k+5)}\gamma_3 \\ & + \frac{3.5(k+2)(k+3)}{2(2h+3)(2h+5)(2h+7)}\gamma_1{}^3\zeta^2\Bigg\} P_h \\ & + \gamma_1\gamma_2\zeta \left\{ \frac{6k(k+2)}{(2k+1)(2h+3)(2h+5)} P_{k-1} \right. \\ & + \frac{3}{2(2h+1)(2k+7)}\left[\frac{(k+1)(2h+5)}{2h+3} + \frac{(k+3)(2h+3)}{2h+5}\right] P_{k+1} \\ & + \frac{4\gamma_3}{(2k+1)(2k+3)}\left\{\frac{k(k-1)}{2k-1}P_{k-2} + \frac{(k+2)^2}{2k+7}P_{k+2}\right\}\Bigg] \end{aligned}$$

and in general

$$\begin{aligned} {}_kf_n = \zeta^{k+n+2}\{ & (\zeta^{n-1}, \zeta^{n-3}\cdots)P_k \\ & + (\zeta^{n-2}, \zeta^{n-4}\cdots)(P_{k-1}, P_{k+1}) + \cdots + (1)(P_{k-n+1}, P_{k+n+1}\}, \end{aligned}$$

where we write

$$(1, \zeta^\sigma \cdots)P_r \equiv (c_{0,r} + c_{\sigma,r}\zeta^\sigma \cdots)P_r$$

denoting by $c_{m,n}$ a specific coefficient. By means of the calculations of the coefficients involved in ${}_kf_1$, ${}_kf_2$ and ${}_kf_3$, harmonics suitable for a number of applications are available in the form

$$v = {}_kf_0 + \lambda{}_kf_1 + \lambda^2{}_kf_2 + \lambda^3{}_kf_3 + O(\lambda^4).$$

It is found that this degree of approximation is adequate for a number of electrical and hydrodynamical problems.

V. *Applications of General Harmonics.*

5.1. It remains now to record the simple manner in which these standard harmonics $\Phi_k(\mu, \zeta)$ can be used to solve the problems mentioned at the beginning of this paper.

All these harmonics are evanescent on the sphere at infinity, since they involve only positive powers of ζ: they are also free from singularities in the region between the surface s and the sphere at infinity. Thus two conditions of the problems are satisfied. The remaining point is to show how harmonics may be obtained which on the surface s, given in our co-ordinates $(\mu\,\zeta)$ as the unit ζ level, take certain specific forms.

Now on the surface s

$$\begin{aligned}\Phi_k(\mu, \zeta) &= \Phi_k(\mu, 1)\\ &= P_k + \lambda(P_k) + \lambda^2(P_k, P_{k\pm1}) \cdots + \lambda^n(P_k, P_{k\pm1}, \cdots, P_{k\pm(n-1)})\end{aligned}$$

where the terms in $\lambda^r P_s$ have specific coefficients. Thus to find a harmonic which reduces to a multiple of P_k on the surface s we take

$$\begin{aligned}v = \Phi_r &+ \lambda^2(\Phi_{k\pm1}) + \lambda^3(\Phi_{k\pm1}, \Phi_{k\pm2})\\ &+ \cdots + \lambda^n(\Phi_{k\pm1}, \Phi_{k\pm2}, \cdots, \Phi_{k\pm(n-1)})\end{aligned}$$

(where again each term has a specific coefficient) choosing the coefficients in such a way as successively to make the P_{k+1}, P_{k-1}, P_{k+2}, $P_{k-2} \cdots$ terms in V disappear, when we successively take

$$\begin{aligned}V &= \Phi_k + O(\lambda^2)\\ V &= \Phi_k + \lambda^2(\Phi_{k\pm1}) + O(\lambda^3)\\ V &= \Phi_k + \lambda^2(\Phi_{k\pm r}) + \lambda^3(\Phi_{k\pm1}, \Phi_{k\pm2})\end{aligned}$$

and so on.

This procedure allows harmonics to be constructed which on the surface s take any specific form of the type

$$C_0P_0(\mu) + C_1P_1(\mu) + C_2P_2(\mu) \cdots + C_kP_k(\mu) + \cdots.$$

Conclusion.

We do not now proceed any further with the study of these standard harmonics for the general surface of revolution under consideration. We have fully investigated their structure. The interesting fact that they form [illegible]

hydrodynamics which we have cited, to any degree of approximation whatever, resides in the heavy arithmetic which may well prove irksome. However, if the applications warrant the trouble involved, there is no reason why the coefficients of these harmonics should not be tabulated, much as the values of functions which are useful in physics have been tabulated. This work could be carried out by arithmeticians with no knowledge of mathematics.

The important point however from our point of view is that these harmonics exist and have a simple structure. This simplicity largely depends upon the fact that each type of surface leads on to harmonics available for rather more general surfaces which are, nevertheless, still merely sets of $P_k(\cos u)$ functions multiplied by powers of e^v. The general determination of the coefficients involved is possibly hardly feasible, though investigations in connexion with them easily yield interesting theorems about the functions.

A Continuum Every Subcontinuum of which Separates the Plane.

By G. T. Whyburn.*

1. *Introduction.* In this paper there will be constructed an example of a continuum of the type, hitherto unknown, indicated in the above title. Indeed, we shall construct a compact plane continuum M having the following properties:

1). *M is the common boundary of two domains;*
2). *every subcontinuum of M separates the plane;*
3). *every subcontinuum of M contains a continuum which is homeomorphic with M, or, in other words, M is topologically contained in each of its subcontinua;*
4). *M contains no uncountable collection of mutually exclusive subcontinua and therefore no indecomposable continuum; obviously it contains no arc;*
5). *M admits of upper semi-continuous decomposition* † *into elements (continua or points) all save a countable number of which are points and with respect to which M is a simple closed curve; clearly, then, all the "point elements" in this decomposition are local separating points of M and are points of ordinary order 2 of M;*
6). *M contains two continua which are not homeomorphic with each other.*

In addition, it will be shown, in general, that any compact plane continuum having property 2) contains a continuum having properties 1) and 2) and that every continuum having 1) and 2) also has properties 4), 5), and 6). From these facts it follows that no plane continuum which separates the

* John Simon Guggenheim Memorial Foundation Fellow.

† That is, an upper semi-continuous collection G of subsets g of M exists such that every point of M belongs to exactly one element of G. A collection G of point sets is upper semi-continuous provided that if g is any element of G and $g_1, g_2, g_3, \cdots$ is any infinite sequence of elements of G containing points $P_1, P_2, P_3, \cdots$; respectively, such that the sequence of points $P_1, P_2, P_3, \cdots$ has a sequential limit point belonging to g, then g contains the entire sequential limiting set of the sequence of sets $g_1, g_2, g_3, \cdots$. See R. L. Moore, "Concerning Upper Semi-continuous Collections of Continua," *Transactions of the American Mathematical Society*, Vol. 27 (1925), pp. 416-428.

plane can be homeomorphic with every one of its subcontinua. So far as the author knows, no example, outside of that of a simple continuous arc, is known of a continuum which is homeomorphic with every one of its subcontinua,* and prior to the example given in this paper, no continuum, other than an arc, is known which is topologically contained in every one of its subcontinua.

To give a general idea of the construction of the continuum M, we list still another property of M, which is stronger than property 5), as follows:

> 7). *M can be decomposed into a contracting † collection G of mutually exclusive elements every one of which is either a point or a continuum of condensation of M which is homeomorphic with M and such that with respect to the elements of G, M is a simple closed curve, and the elements of G which are not points are dense in M.*

We begin, then, with the idea in mind to construct a continuum M having property 7) and also such that a) every element of G which is not a point is a continuum of condensation of every subcontinuum of M of which it is a proper part, and b) if $E_1, E_2, E_3 \cdots$ is a sequence of continua such that E_1 is an element of G and, for $n > 1$, E_n is an element of the collection of elements in E_{n-1} corresponding to the collection G in M, then the diameter of E_n approaches zero as n increases indefinitely. It will be apparent from the construction and proof given below that our continuum M has these properties. And it is true, furthermore, that any compact plane continuum having properties 7), a), and b) must also have properties 1)—6).

Very generally speaking, M is constructed as follows. We take a square Q, which we call the germ cell for M. Take a non-dense perfect set R on Q whose complementary segments are all of length $\leqq 1/9$ the side of Q. Within Q and in the proximity of each complementary segment s of R we introduce a germ cell square $q(s)$ of side 1/9 the length of s, which will be developed exactly like the square Q, and then subject s to a certain *prolongation* process, in such a way that the prolongation of s approaches asymptotically the entire cell which is developed from $q(s)$. The prolongation process is, of course, carried out simultaneously with the development of $q(s)$. At the same time as s is being deformed or prolonged by stages,

* In this connection, see J. R. Kline, "Separation Theorems and Their Relation to Recent Developments in Analysis Situs," *Bulletin of the American Mathematical Society*, Vol. 34 (1928), p. 192.

† A collection of point sets is contracting if, for each $\epsilon > 0$, at most a finite number of elements of the collection are of diameter $> \epsilon$. This terminology is due to R. L. Moore.

the previous prolongations of s are divided into intervals and these developed in the same way as is a side of Q. After this process is carried to completion, the continuum M is defined as the sum of the residue R of Q, the corresponding residues on the deformations of the deformed segments in Q, together with the limit points of these deformations. As so constructed and defined, M will admit of a decomposition of the type in property 7), the elements of G being the cells such as $q(s)$ and corresponding cells resulting from the deformation of segments from the prolongation of s, and from the prolongations of these segments, and so on, together with all points in the residues above mentioned and all of the added limit points not in some cell element.

2. *Development Stages.* We first define some terms which it will be convenient to use for the sake of brevity. A *coupling joint* of bore b and length h is a set congruent to the sum of the two intervals from $(0,0)$ to $(h,0)$ and from $(0,b)$, to (h,b). An ordinary *elbow joint* of bore b and branch lengths h and h' is a set congruent to the sum of the two broken lines joining, in order, $(0,h)$, $(0,0)$, $(h',0)$ and (b,h), (b,b), (h',b) respectively. A *reducing elbow joint* of ratio $b:b'$ is a set congruent to the sum of the two broken lines joining, in order, the points $(0,h)$, $(0,0)$, $(h',0)$ and (b,h), (b,b'), (h',b') respectively. In the construction below, the bores and lengths of the joints used are usually predetermined by the position into which the joints must fit, and hence they are given below only when this is not the case. All reducing elbows used are of ratio $9:1$. A coupling joint is said to be *rectangular* with a segment s when either of its sides may be joined by lines to s so as to form a rectangle.

A *loop* of *type 0* in a segment s of length L is simply a coupling joint of bore L and length L, which is rectangular with s and on the side of s opposite to the deformations of s at a distance L from s. A loop of type 1 in such a segment s is a coupling joint of bore $L/9$ and length L which is rectangular with s and is on the side of s opposite to the deformations of s and at a distance $L/9$ from s.

A segment s is said to be in *stage 0* of deformation when it is chosen for deformation. A segment s is put in stage 1 of deformation [illegible]

(9, 0), (9, 3), (3, 3), (3, 6), (6, 6), (6, 4), respectively. The width or *bore* of the first prolongation of s is $L/9$.

An interval I is put into *stage 1* of *interval development* by dividing it into 9 equal subintervals, these ordered from one end to the other, and the four alternate segments thus formed, beginning with the second, chosen for deformation.

A square (or germ cell) is put into *stage 1* of *cell development* by putting each of its sides into stage 1 of interval development, thus forming a *growing cell in stage 1.*

Now assuming that loop types and deformation and development stages have been defined for all integers $< n$, $(n > 1)$, we proceed to define these notions for n as follows.

A segment s of length L which is in stage $n-1$ of deformation is put into stage n as follows:

a). its central growing cell is put into stage $n-1$;

b). its $(n-1)$st prolongation is divided into a set G_{n-1} of intervals all of length $L/9^{n-1}$, i. e., equal to the bore of this prolongation;

c). for each integer $k \leqq n-1$, the intervals of the set G_k are all put into stage $n-k$ of development, with all deformations of segments placed on the outside of the kth prolongation of s, i. e., on the side opposite to that occupied by the other side of this prolongation;

d). in its germ cell square, in each segment which has been selected for deformation and whose deformation is in the jth stage, $j \leqq n-1$, introduce a loop of type j. Order these loops clockwise, beginning with the lower right vertex of the cell, into a sequence $q_1, q_2, \cdots, q_{4r}$ $(r = 5^{n-1} - 1)$. Now it will be seen from the general definition of loops given below that the bore of each of these loops is $L/9^n$. The bore of the $(n-1)$st prolongation of s being $L/9^{n-1}$, we may join this prolongation of s to q_1 by a reducing elbow with reducing ratio 9:1. For each $i = mr$, $(m = 1, 2, 3)$, join the more advanced end of q_i to the less advanced end of q_{i+1} by an ordinary elbow joint. For every remaining $i < 4r$, join the more advanced end of q_i to the less advanced end of q_{i+1} by a coupling joint. Finally, attach a coupling joint of length $L/9^n$ on to q_{4r}. Add together the $4r$ loops $[q_i]$ and the $4r+1$ joints, and call the set thus obtained the nth prolongation of s and denote it by $p_n(s)$. Clearly $p_n(s)$ is the sum of two broken lines which will be called its two sides, the one nearer to the central cell being called the *near side* and the other the *far side* of $p_n(s)$. The bore of $p_n(s)$ is $L/9^n$.

For $n > 1$, a loop of type n is introduced into a segment s of length L

which is in the nth stage of deformation according to the following procedure. In each of the segments chosen for deformation on the sum of the first $n-1$ prolongations of s whose deformation is in the rth stage, $r \leqq n-1$, introduce a loop of type r. Let these loops be ordered, beginning with the initial point of the near side of $p_1(s)$, which is the right end point of s, and proceeding in order along the near sides of $p_1(s)$, $p_2(s) \cdots$ to the final end of $p_{n-1}(s)$, in a sequence $q_1, q_2, \cdots, q_k$. Then beginning with the final end of the far side of $p_{n-1}(s)$, proceed backwards, in order, along the far sides of $p_{n-1}(s)$, $p_{n-2}(s) \cdots$ to the initial end of the far side of $p_1(s)$, (the right end point of s), with the sequence $q_{k+1}, q_{k+2}, \cdots, q_m$. Now the bore of each of these loops $[q_i]$ is $L/9^n$. For we have assumed that for all integers $r < n$ and each segment s' of length h which is in the rth stage of deformation, a loop of type r in s' is of bore $h/9^r$. And since for each segment s' corresponding to a loop q_i and being in stage r, $h = L/9^{n-r}$, it follows that the bore of the loop q_i corresponding is $1/9^r \cdot L/9^{n-r} = L/9^n$. Now for each i, $1 < i < m$ and $i \neq k$, join q_i to q_{i+1} by a coupling joint or an ordinary elbow according as the segments corresponding are or are not co-linear. Join q_k to q_{k+1} by a joint which is the sum of two congruent elbow joints of branch lengths $h = L/9^n$, $h' = 7L/2 \cdot 9^n$. On to q_i and q_m attach elbow joints of branch lengths $h = 3L/9^n$, $h' = 2L/9^n$, the branch of length h being attached to q_1 and q_m respectively and the joints being turned so that the right end point of s is in the angle of the one attached to q_1 and the left end point of s in the one attached to q_m. Add together these m loops $\cdot [q_i]$ and $m+1$ coupling joints, and the set thus obtained is a loop of type n introduced into s. Its bore, as shown above, is $L/9^n$.

For $n > 1$, an interval I in stage $n-1$ of interval development is put into stage n by putting each of the four segments selected for deformation in stage 1 into stage n of deformation, and by placing each of the five complementary intervals of these segments into stage $n-1$ of interval development.

For $n > 1$, a growing cell C in stage $n-1$ of cell development is put into stage n of cell development by putting each of the four sides of its germ cell square into stage n of interval development.

development of M is selected for deformation, let $p_n(s)$ denote the nth prolongation of s; let $p_n'(s)$ denote the residue of $p_n(s)$, i. e. $p_n(s)$ minus all segments which are selected at some stage from $p_n(s)$ for deformation; let $p(s) = \sum_1^\infty p_n(s)$, the total prolongation of s, and $p'(s) = \sum_1^\infty p_n'(s)$, the residue of $p(s)$. Now for any set X, which may be either a square, a pro-

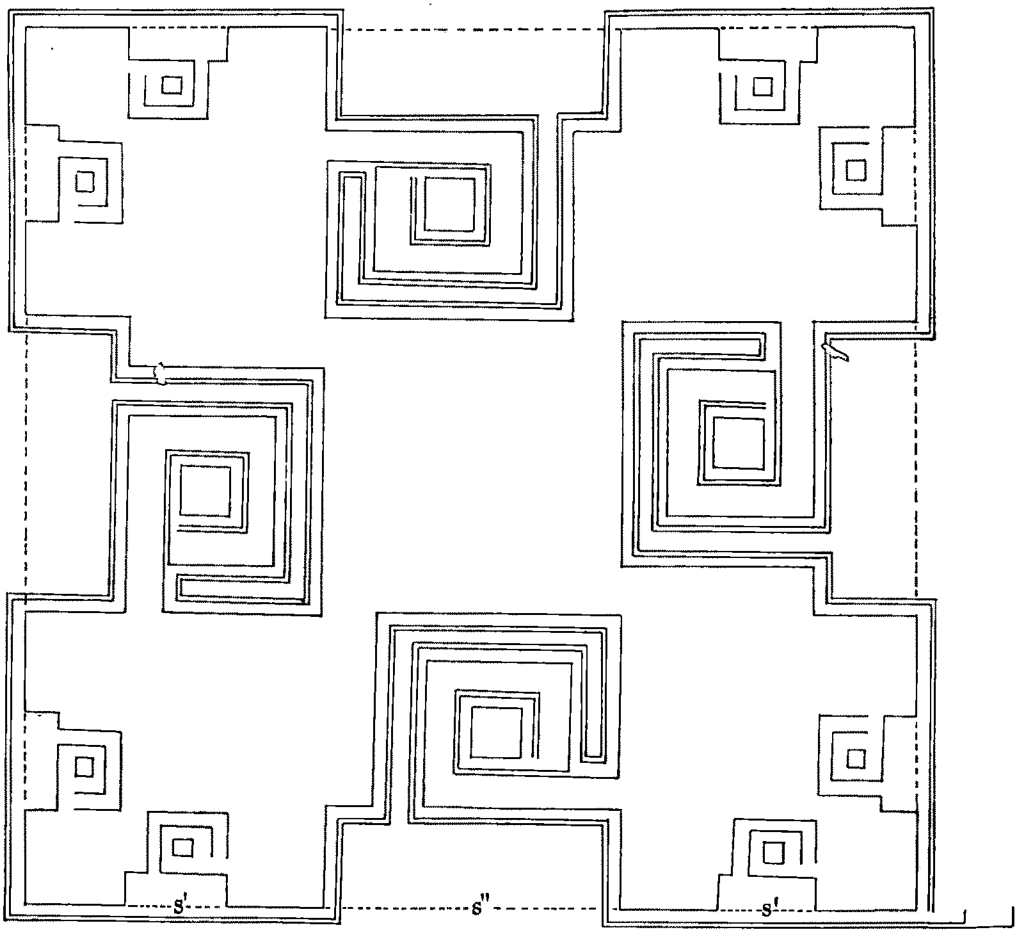


FIG. 1. THE FOURTH PROLONGATION OF A SEGMENT s, WITH CENTRAL GROWING CELL IN STAGE 3 OF DEVELOPMENT. (This figure is not drawn to scale and does not correspond exactly to the description given for the development of M; the sides of the germ cell square for s are divided into 3 parts rather than 9 as in the description.) The segment s is not shown; segments of the type marked s' are in stage 1 of deformation and have corresponding loops of type 1; segments such as the one marked s'' are in stage 2 of deformation and have corresponding loops of type 2.

longation p of some segment, or a collection g of such prolongations, let S denote the operation of selecting a segment from the set X which, at some stage in the development of M, has been chosen for deformation; and denote the segment, which is the result of the operation S or is the value of S, by $S(X)$. For each segment s which, at some stage in the development of M, is chosen for deformation, let $g_1(s) = p(s)$, $g_1'(s) = p'(s)$, and for every $n > 1$, let $g_n(s) = \sum_{S \text{ in } g_{n-1}} p[S(g_{n-1})]$, and $g_n'(s) = \sum_{S \text{ in } g_{n-1}} p'[S(g_{n-1})]$, the residue of $g_n(s)$. Finally, let $d(s) = \sum_1^{\infty} g_n(s)$; $d'(s) = \sum_1^{\infty} g_n'(s)$, the residue of $d(s)$; let R denote the residue of Q, i. e., Q minus all segments selected from Q for deformation, let $F = R + \sum_{s \text{ in } Q} d'(s)$, and let $\bar{F} = M$, the completed cell.

For every square q selected in the development of M as the germ cell of some segment s, it is clear that the completed cell, which we shall denote by $C(s)$, defined with respect to q just as $\bar{F}$ is defined relative to Q, is homeomorphic with M; indeed it is identical in structure with M except in size, and is a subset of M and indeed of $\bar{F} - F$, except when $q = Q$. Let C denote the collection of all completed cells $C(s)$, such that for some segment s selected in Q for deformation, for some n $(1 \leq n < \infty)$, and for some S in $g_n(s)$, $C(s) = C\{S[g_n(s)]\}$, and let T denote the sum of all the point sets of the collection C, i. e.,

$$T = \sum_{s \text{ in } Q} \sum_{n=1}^{\infty} \sum_{S \text{ in } g_n(s)} C\{S[g_n(s)]\}.$$

Now for every element c of C, there exists an integer n and a sequence $s_1, s_2, s_3, \cdots, s_n$ of segments selected for deformation in the development of M, such that (α) s_1 belongs to Q, (β) for every i, $1 < i \leqq n$, s_n is in $p(s_{n-1})$, and (γ) $c = C(s_n)$. Conversely, for every such sequence of segments satisfying (α) and (β), it is clear that $C(s_n)$ is an element of C. Thus C is identically the collection of all cells c completed in the construction of M such that a sequence $s_1, s_2, \cdots, s_n$ of segments selected for deformation exists satisfying conditions (α), (β), and (γ).

Now R is a null dimensional perfect set. And for each segment s [illegible]

$[p(s)]$ is a contracting collection, i. e., at most a finite number of its elements are of diameter $>$ any preassigned positive number. Thus for every such s, $d'(s)$ is a null dimensional F_σ and therefore so is $F = R + \sum_{s \text{ in } Q} d'(s)$.

I shall next show that M is a continuum. By definition it is closed and compact. And from the definition of F it is obvious that if A is the center of Q and B a point without Q, then M is the boundary of both R_a and R_b, the components of the complement of M containing A and B respectively. It is sufficient, then, to show that R_a and R_b are different. If not, then there exists an arc AB from A to B with $AB \cdot M = 0$. Let X_1 be the first point on AB, in the order A, B, of $AB \cdot Q$. Since X_1 does not belong to M and thus not to R, it belongs to a segment s_1 in Q which was selected for deformation in the development of M. Now $Q - s_1 + p(s_1) + C(s_1)$ is a continuum K_1 separating A and X_1. Let X_2 be the first point on AB, in the order A, B, which belongs to this continuum. Then as X_2 is within Q and $C(s_1) \subset M$, X_2 belongs to $p(s_1)$. But since it does not belong to $p'(s_1)$, it belongs to some segments s_2 selected for deformation in the development of M. Now $K_1 - s_2 + p(s_2) + C(s_2)$ is a continuum K_2 separating X_2 and A; let X_3 be the first point of $AB \cdot K_2$ in the order A, B, and so on. Continuing this process indefinitely, we obtain a sequence of mutually exclusive segments $s_1, s_2, \cdots$ which were chosen for deformation in the development of M, and such that for every i, $s_i \cdot AB \neq 0$. But for each i, the end points A_i and B_i of s_i belong to F and hence to M; and since $\lim_{i\to\infty} \delta(s_i) = 0$, therefore every limit point of $\sum X_i$ must belong to M. But this set has at least one limit point P; and since $\sum X_i \subset AB$, it follows that $AB \cdot M \supset P$, contrary to supposition. Therefore R_a and R_b are different, and hence M is a continuum which is the common boundary of the two domains R_a and R_b.

Now by what has been proved above, we may divide M into three sets T, F, and X such that $T + B + X = M$, $T \cdot F = T \cdot X = F \cdot X = 0$, $T + X = \bar{F} - F$. Let G denote the collection whose elements are the elements of C together with all points of $F + X$. Then the elements of G are mutually exclusive, the sum of all of them is M, G is a contracting collection (since C is such a collection), and hence also is upper semi-continuous. I shall now show that every element P of G is a continuum or point of order * two

* A point P of a continuum M is a point of order two of M, in the Menger-Urysohn sense, if for each $\epsilon > 0$, P can be ϵ-separated in M by some two points of M. A point, or continuum, P in M is said to be of order two in M relative to some collection Q of subsets of M if for each $\epsilon > 0$, P can be ϵ-separated in M by two of the elements of Q i. e., two elements A and B of Q exist such that $M - (A + B)$ is the sum of two mutually separated sets M_1 and M_2, where $M_1 \supset P$ and $\delta(M_1 + A + B) < \delta(P) + \epsilon$, [$\delta(x)$ denotes the diameter of the set x].

of M relative both to F and C. If P is an element of C, then there exists a segment s chosen for deformation in the development of M and such that $P = C(s)$. And for any $\epsilon > 0$, an integer k exists so that every point of $\sum_{n=k}^{\infty} p_n(s)$ is at a distance $< \epsilon/2$ from $P = C(s)$. Thus it is clear that if we choose one point of F on each side of $p_k(s)$, or if we choose one segment on each side of $p_k(s)$ which is deformed, then the two points of F so chosen will ϵ-separate P in M and so also will the sum of the two cells corresponding to the two segments thus chosen. If P is a point of F, then either it belongs to R or these exists a segment s which has been chosen for deformation and such that P belongs to $p'(s)$. In either case, for any $\epsilon > 0$, if P is not an end point of any segment chosen for deformation, we can find two points of F, one on either side and belonging to R or to $p'(s)$, respectively, and which ϵ-separate P in M; if P is an end point of some segment s', we can take one of the points on $p_1'(s')$ and the other in R or $p'(s)$ and ϵ-separate P in M, because P is an end point of at most one such segment. In an entirely similar way we can, for each $\epsilon > 0$, ϵ-separate P in M by two elements of C. Finally, if P is a point of X, then since, as is easily seen, $R + \sum_{s \text{ in } Q} \overline{d'(s)} = M$, there exists a segment s_1 in Q such that $P \subset \overline{d'(s_1)}$. Again since, as is easily seen, $\overline{d'(s_1)} = p'(s_1) + C(s_1) + \sum_{s \text{ in } p(s_1)} \overline{d'(s)}$, and P belongs neither to $p'(s_1) \subset F$ nor to $C(s_1) \subset T$, there exists a segment s_2 in $p(s_1)$ such that $P \subset \overline{d'(s_2)}$. Likewise $\overline{d'(s_2)} = p'(s_2) + C(s_2) + \sum_{s \text{ in } p(s_1)} \overline{d'(s)}$, and so on. Continuing this process indefinitely, we obtain a sequence of segments $s_1, s_2, s_3, \cdots$ such that $P \subset \prod_1^{\infty} \overline{d'(s_i)}$. But from the method of development of M, the diameter of $\overline{d'(s_i)}$ approaches zero as i increases indefinitely. Thus $P = \prod_1^{\infty} \overline{d'(s_i)}$; and since for each i, $\overline{d'(s_i)} \cdot \overline{M - \overline{d'(s_i)}} = A_i + B_i$, where A_i and B_i are the end points of s_i, it follows that P is a point of order two of M relative to F. Now if $\epsilon > 0$, there exists an i such that $A_i + B_i$ ϵ-separates P in M; and if we choose two segments e and f, one on each side of $p(s_i)$ such that their sum separates s_i and s_{i+1} in $p(s_i) + s_i$ and such that these segments were chosen for deformation in the development of M, then it is clear that $C(e)$ and $C(f)$ are elements of C and that their sum ϵ-separates P in M.

It has just been shown that every element of G is a continuum or point of order two of M both relative to F and to C. Since M is a compact continuum of elements of the upper semi-continuous collection G of continua and every element of G is an element of order two, it follows that M is a

simple closed curve with respect to the elements of G. Thus since every element of G is either a point or a cell of C which clearly is homeomorphic with M and is a continuum of condensation of M, it follows that M has property 7). Property 1) has already been demonstrated for M, and property 5) is weaker than 7). Therefore M has properties 1), 5), and 7).

4. *Proof of properties* 3), 2), *and* 6) *for* M. We shall first show that M has property 3). To this end we establish the following:

LEMMA. *If a subcontinuum N of M contains a point of some cell E of C and also a point not in E, then N contains E.*

By the definition of the collection C, it follows that there exists a segment s, chosen for deformation in the development of M, such that $E = C(s)$. Now since N contains a point both of $C(s)$ and of $M - C(s)$, it is readily seen that there exists an integer n such that, for every $m > n$, either the near side or the far side, say the near side, of $p_m'(s)$ is a subset of N. But every point of $C(s) = E$ is a limit point of the sum of the residues of the near sides of $p_m(s)$ for $m > n$. Then since N is closed, it must contain E; and the lemma is established.

Now to prove property 3), let N be any subcontinuum of M whatever. Since, as shown in § 3, every point of $F + X$ is a point of order two of M relative to C, it follows that N contains a point in at least one element E_1 of C. And by the lemma, either $N \supset E_1$ or $N \subset E_1$. If the former of these is true, then property 3) is true. If the latter is true, then since E_1 is homeomorphic with M, it follows by the same argument and lemma as above that an element E_2 of the collection in E_1 corresponding to the collection C in M exists such that either $N \supset E_2$ or $N \subset E_2$. If the former, then property 3) is true, because E_2 is homeomorphic with M. If the latter, then an element E_3 in the collection E_2 corresponding to C in M exists such that either $N \supset E_3$ or $N \subset E_3$; and so on. Continue this process. If N contains any of the continua $E_1, E_2, E_3, \cdots$, then since each of these is homeomorphic with M, property 3) hold for M. If not, then $N \subset \prod_1^\infty E_i$. But clearly this is impossible, because $E_i \supset E_{i+1}$, for every i, and $\delta(E_{i+1}) < (1/9)\delta(E_i)$, and therefore $\prod_1^\infty E_i$ contains only one point. Therefore, for some i, $N \supset E_i$, and property 3) is established.

Since M is non-dense and separates the plane, and since every subcontinuum N of M contains a continuum H which is homeomorphic with M, it

follows that H separates the plane and, as N is non-dense, so also does N. Thus M has property 2).

That M has property 6) follows rather obviously from the fact already established that it has property 7). For if t is any simple continuous arc of elements of G and P is an ordinary point which is an interior element of t, then P is a cut point of the continuum K which is the sum of all the elements of G belonging to t. But M has no cut points, and therefore M and K are not homeomorphic with each other. Indeed it is interesting to note that M *contains exactly 4 distinct types of continua.* These are (1) continua homeomorphic with M, (2) continua homeomorphic with arcs of elements of G both of whose end elements are ordinary points, (3) continua homeomorphic with such arcs both of whose end elements are cells of C, and (4) continua homeomorphic with such arcs one end element of which is a point and the other a cell of C. That these are all possible types follows at once with the aid of the lemma in this section. It should be remarked, however, that although every continuum having property 2) contains at least four, one can construct continua having properties 1)—7) and which contain more than four types of continua—indeed containing infinitely many.

Thus we have shown that M has properties 1), 2), 3), 5), 6), and 7). That M also has property 4), although immediately deducible from the construction of M, follows also from the results in the next section concerning continua in general which have properties 1) and 2).

5. *General theorems on continua having property* 2). Let M be any plane continuum having property 2), and let G be any collection of mutually exclusive subcontinua of M. Since each element of G separates the plane, it follows by a theorem of Kuratowski's * that if G is uncountable, it contains a continuum K no proper subset of which separates the plane; but since K contains a proper subcontinuum N, this is contrary to the fact that every subcontinuum of M separates the plane. Hence G must be countable, and thus we have the following theorem:

THEOREM 1. *If every subcontinuum of a plane continuum M separates the plane, then M contains no uncountable collection of mutually exclusive*

boundary of two domains and which, therefore, has properties 1) and 2). Now since, by Theorem 1, K contains no indecomposable continuum, and since no continuum whatever is the sum of a finite or countable number of continua of condensation of itself, it follows that K is not the sum of a finite or countable number of indecomposable continua and continua of condensation of K. Therefore, by results due to Kuratowski,* K may be decomposed upper semi-continuously into a collection G of mutually exclusive elements (continua or points) with respect to which K is a simple closed curve. By Theorem 1, all save a countable number of the elements of G must be points. And clearly each element of G which is a point is local separating point of M and is a point of ordinary order 2 of M. Thus we have the following theorem:

THEOREM 2. *If K is a plane bounded continuum which is the common boundary of two domains and every subcontinuum of K separates the plane, then K admits of upper semi-continuous decomposition into a collection G of elements, (continua or points) all save a countable number of which are points, and with respect to which M is a simple closed curve. All the "point elements" of G are local separating points of K and are points of order 2 of K.*

Now if t is any arc of elements of G, and T is the set of all points of K each of which belongs to some element of G in t, then since, as is evident from Theorem 2, the continuum T has a cut point whereas K does not, then K and T cannot be homeomorphic. Thus every continuum M having property 2) also has property 6). Indeed, we may state the following theorem:

THEOREM 3. *Every plane continuum which separates the plane contains two continua which are not homeomorphic with each other.*

For if this is not so, then M must be homeomorphic with each of its subcontinua. And as M separates the plane, then † every subcontinuum of M must separate the plane, and thus M must have property 2). But then, as shown above, M contains two continua K and T which are not homeomorphic with M, contrary to supposition.

VIENNA, DECEMBER, 1929.

---

* "Sur la structure des frontières communes à deux régions," *Fundamenta Mathematicae*, Vol. 12 (1928), pp. 20-42.

† See Brouwer, *Mathematische Annalen*, Vol. 72 (1912), p. 422.

On Continuous Curves and the Jordan Curve Theorem.†

By LEO ZIPPIN.

We shall show that a continuous curve ‡ which satisfies the Jordan Curve Theorem § non-vacuously ¶ is homeomorphic with the complement on a simple closed surface of a closed and totally disconnected (possibly vacuous) point set. Such surfaces we have previously introduced (J. M. T.) and called cylinder-trees: they represent a generalisation of the simple closed surface and of the number-plane. We shall consider further characterisations of them, equivalent in the presence of the continuous curve spaces to the postulation of the Jordan Curve Theorem, and achieve in particular a considerable reduction of R. L. Moore's axiom-system, $\Sigma 1$, for the number-plane.‖

1. We shall have need of two general theorems.

THEOREM 1. *If a continuum G^* not a continuous curve is a subset of a continuous curve J^*, then G^* has a subcontinuum G which is not a continuous curve such that G is a subset of a maximal cyclic curve J of J^*.*††

We shall content ourselves with an indication of the proof.‡‡ By the

† Presented to the American Mathematical Society, August 30, 1929.

‡ A metric, locally compact, connected, and connected im kleinen space: for definitions of these terms the reader is referred to Hausdorff's "*Mengenlehre,*" 1927.

§ If K is a simple closed curve of a space S, then $S - K$ is the sum of two connected sets without common point such that K is the boundary of each of them.

¶ That is, contains at least one simple closed curve. Compare § 9 of my paper (J. M. T.): "A Study of Continuous Curves and their Relation to the Janiszewski-Mullikin Theorem," *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 744-770.

‖ (F. A.): "On the Foundations of Plane Analysis Situs," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 131-164.

†† For definitions and theorems relating to cyclicly connected continuous curves the reader is referred to the interesting papers of G. T. Whyburn: "Cyclicly Connected Continuous Curves," *Proceedings of the National Academy of Sciences*, Vol. 13 (1927), pp. 31-38 (C. C.) and "Concerning the Structure of a Continuous Curve," *American Journal of Mathematics*, Vol. 50 (1928), pp. 167-194 (C. S.). The theorems of these papers which we shall use are valid in continuous curve spaces, as we have defined them above.

‡‡ Although the author is not acquainted with any previous statement of the above theorem, it seems to him to be implied in any proof of the sufficiency condition of Theorem 28 (C. S.) which is stated without proof.

theorem that a continuum M_n of a continuous curve J^* is a subset of a continuous curve M^*_n of J^* whose upper distance from M_n may be supposed less than any preassigned e, it is possible to replace the continua (M_n) of G^* (of the Moore-Wilder lemma) by independent arcs (L_n) of J^* with the same sequential continuum of condensation M_a, each arc L_n having at least two points in common with the corresponding M_n. By methods of (C. C.) it is possible to show that infinitely many of these arcs have subarcs in common with a maximal cyclic curve J of J^*, and that infinitely many of the continua of (M_n) as well as M_a have a subcontinuum of points on J. The product of a maximal cyclic curve J of J^* and a continuum G^* of J^*, if it consists of more than a single point, is a continuum G. Since G contains subcontinua of infinitely many of the continua of (M_n) as well as of M_a, it is readily shown that G cannot be connected im kleinen.

THEOREM 2. *If B is a compact subset of a continuous curve S such that the maximal connected subsets of B are points or arcs, and no inner point of any arc b of B is a limit point of $B - b$, then there exists an acyclic continuous curve V of S which contains B.*

Let B^* be the set of points of B which are either components of B or are endpoints of arcs of B. Then B^* is closed and totally disconnected since no inner point of an arc of B can be a limit point of B^*. There is an acyclic continuous curve V^* of S which contains B^*.† Then $T = V^* + B$ is closed and connected, since V^* has at least one point in common with each component of B. Suppose that any subcontinuum G of T fails to be connected im kleinen: then G contains a continuum M of points at which it fails to be connected im kleinen. Suppose that any point m of M fails to be a point of B. Then in some neighborhood of m, G is locally identical with a subcontinuum of V^* and necessarily connected im kleinen at m, since V^* is an acyclic continuous curve. If, on the other hand, every point of M is a point of B, then M being a continuum of points of B is a subarc of B. Let m' be an inner point of the arc M. There is a neighborhood U_m' of m' such that $B \cdot U_m' \subset M$. If m' fails to be a point of V^*, then in some neighborhood of m', G is locally identical with M and therefore connected im kleinen at m'. If m' is a point of V^*, let $U^*_{m'}$ be a neighborhood of m' such that any point of $M \cdot U^*_{m'}$ can be joined to m' by a subarc of $M \cdot U_{m'}$. Let $U''_{m'}$ be a neighborhood of m' such that any point of $V^* \cdot U''_{m'}$ can be joined to m' by an arc of $V^* \cdot U^*_{m'}$. It is readily seen that any point of

† (J. M. T.) Theorem 1.

$G \cdot U''_{m'}$ can be joined to m' by an arc of $G \cdot U_{m'}$. Then G is a continuous curve. Let V be a continuum of T irreducible about B. Then V is a continuous curve. If V contains any simple closed curve K, a point q of K is not a limit point of B and there is an arc Q of K no point of which belongs to B. Since the arc Q contains at least one non-cutpoint of V, it readily follows that V is not irreducibly connected about B. Then V is an acyclic continuous curve.

2. We assume that a continuous curve S has the following properties:

(A) *S contains at least one simple closed curve.*

(B) *if L is an arc of a simple closed curve of S, then $S - L$ is connected.*

(C) *if K is a simple closed curve of S, then $S - K$ is not connected.*

Since S contains at least one simple closed curve it contains a maximal cyclic curve J, and if S is not J then there exists on J a point x which is a cutpoint of S. Then x belongs to a simple closed curve of J and this contains an arc L containing x such that $S - L$ is not connected. Therefore S must have the further property:

(D) S is cyclicly connected. We have shown previously † that (B, C, D) imply that a continuous curve has the property:

(E) *If K is a simple closed curve of S, $S - K$ is the sum of two maximally connected sets (K-domains) such that K is the complete boundary of each of these.* And, further, the property:

(F) *If K is a simple closed curve of S and M is either of its K-domains (domains relative to K), and $\langle amb \rangle$ ‡ is an open arc of M whose endpoints are on K, then $M - \langle amb \rangle$ is the sum of two maximally connected sets which are domains relative to $K + \langle amb \rangle$, such that one of these has the simple closed curve ao_1bma for its boundary and the other has for its boundary the simple closed curve ao_2bma, where o_1 and o_2 are any points of K separating a and b.* It will be remarked that (A, E) imply (B, C) so that (A, E) and (A, B, C) are equivalent: § therefore the spaces which we are considering may properly be termed Jordan Curve Theorem spaces.

† [illegible]

‡ [illegible]

§ [illegible]

2.1. We shall show that if J^* is a continuous curve of S, D'' an open connected subset of $S - J^*$, and G^* a continuum of $J^* \cdot \overline{D''}$, then G^* is a continuous curve.

If D'' is not a maximally connected subset of $S - J^*$ it is contained in a J^*-domain D^*, and it is seen that $G^* \subset J^* \cdot \overline{D^*}$. If G^* is not a continuous curve, by Theorem 1 it contains a continuum G not a continuous curve such that G is a subset of a maximal cyclic curve J of J^*. Since J is a subset of J^*, $S - J$ contains a maximal connected subset D such that $D^* \subset D$; and further $G \subset J \cdot \overline{D}$. If every pair of points of G disconnects G, G is a simple closed curve. Then we may suppose that there is at least one pair of points x and y of G such that $G - (x + y)$ is connected. There is in J a simple closed curve K which contains x and y. We shall show that any other point z of G separates G between x and y. If z does not belong to the simple closed curve K, it is a point of a K-domain, M. Since D is a J-domain and z is a limit point of D, $D \subset M$. Since no point of J is a cut point, there is an arc hzk such that $\langle hzk \rangle \subset M$ and $h + k \subset K$. If h and k separate x and y on K then, by (F), $\langle hzk \rangle$ divides M into two $(k + hzk)$-domains such that one of these has the boundary $hzkxh$ and the other the boundary $hzkyh$. Then since these are simple closed curves of J, D must lie entirely within one of the subdomains of M, and either x or y fails to be a limit point of D. Then h and k cannot separate x and y on K, and there is a simple closed curve $K' = xzyz'x$ (where z' is an arbitrary point on the other xy-arc of K') of J.† Let N be that K'-domain which contains D. If any point of J is not a point of $K' + N$, $J - K'$ cannot be connected and $N \cdot J$ is an open subset of $J - K'$. Then $N \cdot J$ is the sum of components of $J - K'$ (or it is vacuous in which case G is a sub continuum of K' and therefore a continuous curve) and $J' = K' + N \cdot J$ is a continuous curve.‡ Let u be any point of $N \cdot J$ and v a point of J not in $K' + N \cdot J$. There is in J a simple closed curve (uv) and therefore an arc huk such that $\langle huk \rangle \subset N$ and $h + k \subset K'$, h and k being distinct. Then, since no point of K' can separate any two points of K' nor separate a point of K' from a point of $N \cdot J$, it is seen that J' is cyclicly connected.§ Suppose $J' - (x + y)$ is connected: then it contains an arc zz', and this has a last point h' on $\langle xzy \rangle$ and a first point k' thereafter on $\langle xz'y \rangle$ and an open subarc $\langle h'k' \rangle$ of N. Since h' and k' separate x and y on K' it is seen (by the argument above for h and k) that either x or y fails to be a limit point

† If $z \subset K$, then $K' = K$.

‡ See Remark to lemma of (J. M. T.).

§ If $J \subset K' + N \cdot J$, $J' = J$.

of D. Then it follows that $J' - (x+y)$ cannot be connected. Since G is a continuum of J', and $G - (x+y)$ is connected, $G - (x+y)$ belongs to a single component of $J' - (x+y)$ and this component plus x and y is a continuous curve J'' such that $J'' - (x+y)$ is connected. If J'' contains any point of $\langle xz'y \rangle$ we are led to contradiction by the argument on h' and k'. If the point z does not separate x and y in J'', there is an arc xy of $J'' - z$. This arc has a last point m on $xz - z$ (of K') and a first point n thereafter on $yz - z$. The subarc mn having its endpoints only on xzy, and no point in common with $\langle xz'y \rangle$ since $(\langle xz'y \rangle) \cdot J'' = 0$, has its endpoints only in common with K'. Since $mn \subset J'$, $\langle mn \rangle \subset N$. If the point m is not the point x, the points m and n separate x and z on K', and (as above) either x or z fails to be a limit point of D. Similarly, the point n cannot fail to be the point y. Then mn is an arc xy, such that $\langle xy \rangle \subset N$. Where z'' is an arbitrary point of this arc, there is a simple closed curve $xz''yzx$ of J''. But it can be shown that $x+y$ separates z and z'' in J'' precisely as it was shown that $x+y$ separated z and z' in J'. Since $J'' - (x+y)$ is connected it follows that z must separate x and y in J''. Since $J'' \supset G$, every point of G other than x and y must separate x and y in G, and G is an arc. Then, in every event, G is a continuous curve, and therefore G^* is a continuous curve.

3. We shall prove that if C is the sum of a finite number of disjoint continuous curves of S and x is any point of C which is on the boundary of a domain D relative to C, then x is accessible from D. For if x is not accessible from D it is readily shown † that there exists an ϵ and a sequence (x_n) of points converging to x, such that each point x_n is contained in $D \cdot U_{x\epsilon}$ ‡ while no two points of this sequence can be joined by an arc of $D \cdot U_{x\epsilon}$. Then there exists an infinite set of arcs $(x_n y_n)$ no two of which have any common point, such that each is a subarc of D, and each arc except for its y_n-point belongs to $U_{x\epsilon}$. The set of arcs has a subsequence $(x'_n y'_n)$ with a sequential continuum of condensation T. From the connectivity im kleinen of D it follows that if a point of $T \cdot U_{x\epsilon}$ belongs to D two distinct points x'_j and x'_k are on a subarc of $D \cdot U_{x\epsilon}$. Then $T \cdot U_{x\epsilon} \subset C \cdot \overline{D}$ and since $T \subset \bar{U}_{x\epsilon}$ while C is closed, $T \subset C$. Then T belongs to one of the continuous curves of C: call this F; then T is a continuous curve (§ 2. 1) and contains an arc xy such that $xy - y \subset U_{x\epsilon}$. Let p and q be points of xy distinct from

† Compare R. L. Wilder: "Concerning Continuous Curves," *Fundamenta Mathematicae*, Vol. 7 (1925), Theorem 1, pp. 312-315.

‡ $U_{x\epsilon}$ is a neighborhood of x of diameter less than ϵ.

each other and from x and y. Let U_p and U_q be neighborhoods of p and q respectively such that $\overline{U}_p + \overline{U}_q \subset U_{x\epsilon}$, $\overline{U}_p \cdot U_q + U_p \cdot \overline{U}_q = 0$, and $(U_p + U_q) \cdot (C - F) = 0$. By the connectivity im kleinen of F, there exist neighborhoods U^*_p and U^*_q such that any point of $F \cdot U^*_p$ and $F \cdot U^*_q$ respectively can be joined to p and q respectively by an arc of $F \cdot U_p$ and $F \cdot U_q$ respectively. There exist neighborhoods U''_p and U''_q such that any point of U''_p and U''_q can be joined to p and q respectively by arcs of U^*_p and U^*_q respectively. Since p and q are points of a sequential continu of condensation of $(x'_n y'_n)$, there is an arc $x'_k y'_k$ which has a point p' in U''_p and a point q' in U''_q. There is an arc $p'p$ of U^*_p and this has a last point d on $x'_k y'_k$ and a first point d' thereafter on F; the subarc dd' belongs, clearly, to $D \cdot U_{x\epsilon}$. There is an arc $d'p$ of $F \cdot U_p$: this has a first point a on xy. Similarly there is an arc $cc'b$, where $cc' - c' \subset U^*_q \cdot D$, c' being a point of F and c its only point on $x'_k y'_k$, and $c'b \subset F \cdot U_q$, b being its only point on xy. There is a subarc ab of xy, and a subarc cd of $x'_k y'_k$. Then $ab + bc' + c'c + cd + dd' + d'a$ is a simple closed curve K such that $\langle c'cdd' \rangle \subset D \cdot U_{x\epsilon}$, and $c'badd' \subset F \cdot U_{x\epsilon}$. If any arc $x'_j y'_j$ $(j \neq k)$ has a point in common with each of the K-domains, it has a point in common with K and this is necessarily a point of $\langle c'cdd' \rangle$: then there is an arc $x'_j x'_k$ of $D \cdot U_{x\epsilon}$. Since this contradicts the construction of the original sequence (x_n), every arc of $(x'_n y'_n)$, $(n \neq k)$, belongs to one or the other of the K-domains: let M be a K-domain which contains an infinite set $(x''_n y''_n)$ of the arcs of $(x'_n y'_n)$. Choose distinct points p'', p^*, q^*, q'', on $\langle ab \rangle$ in order: $ap''p^*q^*q''b$. It can be seen that there exist neighborhoods U_{p^*} and U_{q^*} such that:

$$U_{p^*} \cdot [p''adcbq^* + (C - F) + (S - U_{x\epsilon}) + \overline{U}_{q^*}] = 0$$
$$= U_{q^*} \cdot [q''bcdap^* + (C - F) + (S - U_{x\epsilon}) + \overline{U}_{p^*}].$$

Find neighborhoods U''_{p^*} and U''_{q^*} as above for p and q. Then we can construct an arc $a'd^*d''c''c^*b'$ such that a' and b' are on ab in the order: $ap''a'b'q''b$; $\langle a'd^*d''c''c^*b' \rangle \subset M \cdot U_{x\epsilon}$, $(a'd^* - d^*) + (b'c^* - c^*) \subset F$, $\langle d^*d''c''c^* \rangle \subset D \cdot U_{x\epsilon}$, and $d''c''$ belongs to an arc $x''_m y''_m$. Let a'' be any point of $\langle a'b' \rangle$. Since a'' as well as p'' are limit points of the sequential set $(x''_n y''_n)$ it is seen that there is an arc $x''_j y''_j$ $(j \neq m)$ which has a point in each of the subdomains of M determined by the arc $\langle a'd''c''b' \rangle$, since a'' and p'' separate a' and b' on K: property (F). Then $x''_j y''_j$ has a point in common with the arc $a'd''c''b'$ and this is necessarily a point of $d^*d''c''c^*$: there is an arc $x''_m x''_j$ in $D \cdot U_{x\epsilon}$. This contradiction of our construction of the sequence (x_n) proves that x is accessible from D.

3.1. We shall show that no acyclic continuous curve T of S can dis-

connect S. Let T be an acyclic continuous curve of S and D_x and D_y two distinct T-domains. Then T contains an acyclic continuous curve T' irreducible about $G = T \cdot \overline{D}_x$: the endpoints of T' are a subset of G. It is seen that D_x is a T'-domain. Let D'_y be that T'-domain which contains D_y. Then T' contains an acyclic continuous curve T'' irreducible about $G' = T' \cdot D'_y$: the endpoints of T'' are a subset of G'. Since G and G' separate S, they contain at least two points. Let p and q be two endpoints of T''. Then they are points of T', and being limit points of D'_y which is a T'-domain they are accessible from D'_y. There is an arc $py'q$, $\langle py'q \rangle \subset D'_y$. Let $p'pqq'$ be the *maximal* arc of T' which contains pq: p' and p, as well as q' and q, are not necessarily distinct. Since p' and q' are endpoints of T' and therefore points of G, they are accessible from D_x. There is an arc $p'x'q'$, $\langle p'x'q' \rangle \subset D_x$. Let M be that domain of $K = p'x'q'qy'pp'$ which contains $\langle pq \rangle$ of T''. Then $M - \langle pq \rangle = M_1 + M_2$, where M_1 and M_2 are $(K + pq)$-domains and the boundary of M_1 is $py'qp$ while the boundary of M_2 is $pp'x'q'q$. Since the point x' of D_x does not belong to M_1 and the boundary of M_1 is a subset of $T' + D'_y$, it is seen that M_1 can contain no point of D_x. Then M_1 can contain no point of G. Suppose that M_1 contains a point z of T'. Let z^*pzz' be any maximal arc of T' which contains z and p. This arc in order from z to p has a first point z'' on pq. If the arc $z''zz'$ has any point other than z'' on pq it is seen that T' contains a simple closed curve. Then $z'zz'' - z'' \subset M_1$, since $(z'zz'' - z'') \cdot py'q = 0$; and $z' \subset M_1$: but z' is an endpoint of T' and therefore a point of G. Then M_1 contains no point of T' and therefore no point of T''. Similarly, M_2 can contain no point of D'_y and therefore no point of T''. Then $T'' \cdot (K + M) = pq$. If T'' is not the arc pq, T'' has a point k in N, the K-domain distinct from M. Let kp be the arc of T'' containing k and p: p or q is the first point which this arc has on K, and one of these points fails to be an endpoint of T''. Then T'' is the arc pq. But $S - T''$ is not connected, while the arc pq being a subset of a simple closed curve of S cannot disconnect S. The contradiction shows that $S - T$ must be connected.

3.2. Let C be a finite set of disjoint continuous curves of S such that D_x and D_y are distinct domains relative to C. Let $G = C \cdot \overline{D}_x$, and let $G' = G \cdot D$ where D is that G-domain which contains D_y. [illegible]

[illegible]

as in § 3, the arc $ad'dcc'b$ and the simple closed curve K. Here, however, points of $c'b - b$ and of $d'a - a$ although they belong to F may also be points of D. Since a and b are points of G' they are accessible from D_x: there is an open arc $\langle azb \rangle$ of D_x. Let H be that domain of the simple closed curve $azbc'cdd'a$ which contains $\langle ab \rangle$ of G'. Then $H - \langle ab \rangle$ is the sum of two domains M and N with the boundaries $abc'cdd'a$ and $azba$ respectively. It is seen that no point of N can be a point of D, and therefore no point of N can be a point of G' (compare § 3.1). Similarly no point of M can be a point of G'. Then $G' \cdot H = G' \cdot (M + \langle ab \rangle + N) = \langle ab \rangle$. We can continue the argument of § 3, precisely as there, with the added condition on our neighborhoods U_{p^*} and U_{q^*} that they are subsets of H, since now if c^* and d^* are taken as the first points of G' on the respective arcs $c''q^*$ and $d''p^*$ they must be points of $\langle ab \rangle$, and the open arc $\langle a'd''c''b' \rangle \subset D$.

Let fgh be any arc of G'. Since f and h are accessible from D_x and from D, there is an open arc $\langle fxh \rangle$ of D_x and an open arc $\langle fyh \rangle$ of D: let K'' be the simple closed curve $fxhyf$, and let M'' be that K''-domain which contains the arc $\langle fgh \rangle$. By methods we have several times used, it follows that no point of $\langle fgh \rangle$ is a limit point of $G' - fgh$. If then G' contains no simple closed curve, the maximal connected subsets of G' are points or arcs such that if g' is an arc of G' no inner point of g' is a limit point of $G' - g'$: that is, G' is a set B described in Theorem 2. Then G' is a subset of an acyclic continuous curve V of S which is irreducibly connected about it. If x and y are points of D_x and D respectively, G' separates S between x and y. Since x and y belong to some simple closed curve of S, it is readily inferred from property (E) that $x + y$ cannot disconnect S. Moreover x and y are not points of G'. Then there is an acyclic continuous curve V' of $S - (x + y)$ which contains G' † and, accordingly, separates S between x and y. But by § 3.1 no acyclic continuous curve of S can disconnect S. Therefore G' must contain at least one simple closed curve J.

Let $J \equiv acbda$, and let M be that J-domain which contains D. Since every point of G' is accessible from D there is an arc $ay'b$, $\langle ay'b \rangle \subset M$. If D_x has any point in M, $D_x \subset M$. In this case it is readily seen that D_x belongs to one of the subdomains of M determined by the arc $\langle ay'b \rangle$ and that either c or d fails to be a limit point of D_x. Since this is not possible, $D_x \cdot M = 0$, and J separates S between D_x and D. If N is the J-domain which contains D_x, it is seen that $N \cdot D = 0$, and it follows that $G' = J$.

---

† By the argument of (J. M. T.), Theorem 2, the second paragraph: apply this argument to x and y, replacing $C - x$ by $S - (x + y)$.

3.3. If x and y are two points of S, and ϵ is any preassigned positive number (we may suppose ϵ less than the distance of x to y), there is a finite set C of continuous curves of S, ϵ-separating x.† Then there are two distinct C-domains, D_x and D_y. Then it follows, by § 3.2, that there is a simple closed curve J of C which separates x and y in S. No point of J is at a distance from x greater than ϵ, since $J \subset C$. Let H be that J-domain which contains x. There is an ϵ_1 such that any point of S whose distance from x is not greater than ϵ_1 can be joined to x by an arc of H, and such that ϵ_1 is less than $\epsilon/2$. Let J_1 be a simple closed curve which separates x and y, such that no point of J_1 is at a distance from x greater than ϵ_1, and let H_1 be that J_1-domain which contains x. It is seen that $H \supset J_1 + H_1$. We can construct, inductively, a sequence of simple closed curves (J_n) converging to x such that each curve of (J_n) separates x and y and such that if H_n is that J_n-domain which contains x, then for every m, $H_m \supset J_{m+1} + H_{m+1}$. Then $\prod_{n=1}^{\infty} H_n \supset x$; if $\prod_{n=1}^{\infty} H_n \supset x'$ distinct from x, it is seen that every arc $x'y$ of S must have at least one point in common with each of the simple closed curves of (J_n) and therefore it must contain the point x to which these curves converge. Then x separates x' and y in S: but no point of S is a cutpoint. Then $\prod_{n=1}^{\infty} H_n = x$, and for any preassigned ϵ' there is a domain H_m containing x of diameter less than ϵ': J_m ϵ'-separates x in S.

It is now possible to show that S satisfies the Axioms 1-4 of R. L. Moore's $\Sigma 1$ (F. A.) as we have shown it in (J. M. T.) §§ 8 and 8.1 if S is compact and in the last paragraph of § 11 if S is locally compact. The Axioms 6 and 7 are established in (J. M. T.) on the basis of the Janiszewski-Mullikin Theorem. From the considerations of the preceding paragraphs of this paper we shall show that they are consequences of our assumptions (A, B, C). Let x and x' be any two points of a simple closed curve J of S, and let J' be a simple closed curve of S ϵ-separating x, where ϵ may be supposed less than the distance of x to x', and arbitrarily small greater than zero. Since x is accessible from M, where M is either of the J-domains in S, there is an arc px such that $px - x$ belongs to M. This has a subarc qx such that $qx - x$ has no point in common with $J' + J$. Let $C = J' + J$: $qx - x$ belongs to a C-domain D. Then D is a subset of a J-domain and, containing q, it belongs to M. Let x'' be any point of that J'-domain which contains x'. Since D has x for a limit point, D belongs to that J'-domain which contains x and since this does not contain x', x'' cannot belong to D.

† (J. M. T.) § 7.

Let $G = C \cdot \overline{D}$: then G contains a simple closed curve J'' separating q from x'' (by § 3. 2). Let D' be that J''-domain which contains q: $D' \supset D$. If any point of J belongs to D', there is an open arc of points of J in D' with distinct endpoints on J'' and this arc divides D' into two subdomains such that D can belong to at most one of them: but then there are points of J'' which fail to be limit points of D. Then, since x is a point of $\overline{D}$, $x \subset J''$. In a sufficiently small neighborhood of x, C is locally identical with J: since $J'' \subset C$, there is on J'' an arc axb of J. There is an arc ab of $J'' - x$: since $J'' \subset G \subset \overline{D} \subset \overline{M} = M + J$, $ab \subset M + J$. Since x' is not a point of $\overline{D}$, x' is not a point of J''. Then a and b separate x and x' on J, and the arc ab of $J'' - x$ is an arc of $(J + M) - (x + x')$: then it has a subarc $a'b'$ such that a' and b' are on J separating x and x', and $\langle a'b' \rangle \subset M$. Since $\langle a'b' \rangle \subset C$, and $\langle a'b' \rangle \cdot J = 0$, $\langle a'b' \rangle \subset J'$ and no point of $\langle a'b' \rangle$ is at a distance from x greater than ϵ. Then we can construct the domains of § 7. 2 (J. M. T.) and it is a ready consequence of the existence of these domains that S satisfies Axioms 6 and 7 of $\Sigma 1$.

Then by § 8. 2 of J. M. T. it follows that S if it is compact is a simple closed surface, and from the last paragraph of § 11 and the succeeding sections of Theorem 6 it follows that S if it is not compact is homeomorphic with the complement on a simple closed surface of a closed and totally disconnected point set. We shall show, in a later paper, that the use of inversion in Theorem 6 is not essential, so that the restriction that the continuous curves there treated be imbedded in Euclidean space of finite dimension can be removed: the notions of compactness and non-compactness of this paper may replace, accordingly, the use of closed and boundedness and of closed and unboundedness of (J. M. T.). It may be remarked that the number-plane is characterised among the cylinder-trees (spaces homeomorphic with the complement on a simple closed surface of a closed and totally disconnected, possibly vacuous, point set) by its non-compactness and the property that every simple closed curve determines in it one compact domain: to characterise the number-plane, therefore, it is not necessary even in (J. M. T.) to resort to inversion, as may be seen from the last paragraph of § 11.

Then we have established the following theorems:

THEOREM 3. *In order that a continuous curve S be a cylinder-tree it is sufficient that it satisfy non-vacuously the Jordan Curve Theorem.*

THEOREM 3'. *In order that a continuous curve be a simple closed sur-*

face it is sufficient that it satisfy non-vacuously the Jordan Curve Theorem, and be compact.

THEOREM 3″. *In order that a continuous curve be a number-plane it is sufficient that it satisfy the Jordan Curve Theorem non-vacuously, that it be not compact, and that every simple closed subcurve determine in it a compact domain.*

That these conditions are moreover necessary is obvious for Theorems 3′ and 3″, and a consequence for Theorem 3 of the fact that no closed and totally disconnected subset of the domains complementary to a simple closed curve on a simple closed surface can disconnect those domains.

In these theorems, further, it is possible to replace the postulation of the Jordan Curve Theorem by the equivalent system (A, B, C) or by the system (B', C) where B' is the property that if L is an arc of S, then $S - L$ is connected. It is readily seen that (B', C) implies (A, B, C).†

4. We shall show that a continuous curve S which has the property (D) of cyclic connectivity, and also has the property: (G) if K is a simple closed curve of S, every point of K is a limit point of $S - K$, and $S - K$ is the sum of two connected sets without common point; satisfies the Jordan Curve Theorem.

Let K be any simple closed curve of S. Then $S - K = M + N$, where M and N are distinct K-domains. If the Jordan Curve Theorem is not satisfied there is a point y of K which fails to be a limit point of one of the K-domains, say N. By (D), it is clear that N has at least two distinct limit points on K. Then there is an arc xyz of K such that $x + z \subset \bar{N}$, but $\langle xyz \rangle \cdot \bar{N} = 0$. By (G), $\langle xyz \rangle \subset \bar{M}$.

4.1. We consider first the case that $\langle xy'z \rangle \cdot \bar{M} = 0$, where $xy'z$ is the other xz-arc of K: $xy'z \subset \bar{N}$. For any point y'' of $xy'z$ it is possible to construct an arc pq, $\langle pq \rangle \subset N$, $p + q \subset xpy''qz \equiv xy'z$, from which it follows that no point of $\bar{N}$ separates x and z in $\bar{N}$. Since $\bar{N}$ is a continuous curve, there is a simple closed curve $K' = xy_1zy_2x \subset \bar{N}$. Let $S - K' = H + D$ in maximal connected sets: since $M \cdot K' = 0$, M belongs to one of the K' [illegible]
[illegible]
[illegible]

It is seen that $\bar{D}-(x+z)$ is connected: it contains an arc y_1y_2 with a last point, which we may suppose to be y_1, on $\langle xy_1z\rangle$ and a first point y_2 thereafter on $\langle xy_2z\rangle$: it is understood that we are not here supposing accessibility, the points y_1 and y_2 having been in no previous way particularized. Then $\langle y_1y_2\rangle \subset D$. From the relation $S-zyxy_2y_1x=M+xy_1+y_2z+(D-\langle y_1y_2\rangle)$ it is seen that the bracketed point set must be connected: else the simple closed curve $zyxy_2y_1x$ fails to disconnect S into not more than two domains. There is an arc pq, $p\subset\langle xy_1\rangle$, $q\subset\langle zy_2\rangle$, $\langle pq\rangle\subset D-\langle y_1y_2\rangle$. From the relation $S-zy_1y_2z=H+\langle y_1xy_2\rangle+(D-\langle y_1y_2\rangle)$, it is seen that $D-\langle y_1y_2\rangle$ cannot be connected and has a maximally connected subset A_z which has no limit point in $\langle y_1xy_2\rangle$: otherwise zy_1y_2z fails to separate S. If any other component (maximal connected subset) of $D-\langle y_1y_2\rangle$ fails to have a limit point in y_1xy_2, then $S-zy_1y_2z$ has at least three components. If A_z has no limit point in $\langle y_1z\rangle$, then $S-zy_2y_1xyz$ has at least three components: one of these is A_z, the other is M, and a third exists containing for example $\langle xy_2\rangle$. From $S-xy_1y_2x=H+\langle y_1zy_2\rangle+(D-\langle y_1y_2\rangle)$ it is seen that $D-\langle y_1y_2\rangle$ has a component A_x such that $\bar{A}_x\cdot\langle y_1zy_2\rangle=0$, and such that this is not true for any other component: then $A_x\neq A_z$. If $\bar{A}_x\cdot y_2x=0$, $S-xy_1y_2zyx$ has at least three components.

Since $(pq)\cdot(\langle y_1xy_2\rangle)\neq 0\neq(\langle y_1zy_2\rangle)\cdot(pq)$, it follows that $\langle pq\rangle\subset A'$, a component of $D-\langle y_1y_2\rangle$ distinct from A_x and from A_z. There is in D an arc joining a point of $\langle pq\rangle$ to a point of $\langle y_1y_2\rangle$: this has a last point h on $\langle pq\rangle$ and a first point k thereafter on $\langle y_1y_2\rangle$. It is seen that $hk-k\subset A'$. Then $\bar{A}_x\cdot\langle y_1k\rangle\neq 0$; otherwise $S-xy_1zqhky_2x$ has at least three components (one of these is M, and another $\bar{A}_x$: a third contains, for example, $\langle y_2q\rangle$). If $\bar{A}_z\cdot\langle y_2k\rangle=0$, $S-zy_1khpxy_2z$ has at least three components. Knowing that $\bar{A}_x\cdot\langle y_2x\rangle\neq 0\neq\bar{A}_z\cdot\langle y_1z\rangle$, we can deduce the existence of an arc rs, $\langle rs\rangle\subset A_x$, $r\subset\langle y_2x\rangle$, $s\subset\langle y_1k\rangle$, and an arc $r's'$, $\langle r's'\rangle\bar{A}_z$, $r'\subset\langle y_1z\rangle$, $s'\subset\langle y_2k\rangle$.

Then $\bar{A}_x\cdot\langle y_2s'\rangle\neq 0$, or $S-s'y_1xy_2zr's'$ has at least three components, and $\bar{A}_z\cdot\langle sk\rangle\neq 0$, or $S-sy_1zy_2khpxrs$ has at least three components. Then there exist two arcs ad and bc such that $\langle ad\rangle\subset A_x$ and $\langle bc\rangle\subset A_z$, and a, b, c, and d are points on $\langle y_1y_2\rangle$ in the order: $y_1abkcdy_2$. There is an arc $r''x'$, $\langle r''x'\rangle\subset A_x$, $r''\subset\langle y_2x\rangle$, and $x'\subset\langle ad\rangle$. There is an arc r^*z', $\langle r^*z'\rangle\subset A_z$, $r^*\subset\langle y_1z\rangle$, and $z'\subset\langle cb\rangle$. Let K'' be the simple closed curve $ax'dcz'ba$ and let U be that K''-domain which contains H: $U\supset H+K'+\langle y_1a\rangle+\langle bkc\rangle+\langle dy_2\rangle$. If V is the other K''-domain, $V\subset D-y_1y_2$. If $\bar{V}\cdot(ax'd+cz'b)=0$, $\bar{V}\cdot K''\subset(ab+cd)$. Then $S-bax'dcz'r^*zy_2xphkb$ has at least three components. Since V contains no point of y_1y_2 it cannot

contain points of both A_x and A_z, and in consequence a point of A_x and of A_z cannot both be limit points of V. Then from the relation above it follows also that no point of $\langle ax'b \rangle$ can be a limit point of V. But in this case $S - abz'cdx'r''xyzy_1a$ has at least three components.

Then we have shown that for no simple closed curve K of S can one of the K-domains have at most two limit points on K, and further that if there is an arc xyz of K such that $\langle xyz \rangle \cdot \bar{N} = 0$, $x + z \subset \bar{N}$, then at least one point y' on the other xz-arc of K is in M, where M and N are the respective K-domains.

4. 2. Returning to § 4, we can say that $\langle xy'z \rangle \cdot \bar{M} \supset y'$, and that $\langle xy'z \rangle$ contains moreover a point y^* of $\bar{N}$. Let us suppose that both x and z are accessible from N. In this case there is an arc $xy''z$, $\langle xy''z \rangle \subset N$. From the relation $S - xy'zy''z = M + \langle xyz \rangle + (N - \langle xy''z \rangle)$, it is seen that $(N - \langle xy''z \rangle)$ must be connected: for $M + \langle xyz \rangle$ is connected and no point of it is a limit point of $(N - \langle xy''z \rangle)$, nor is any point of $(N - \langle xy''z \rangle)$ a limit point of $M + \langle xyz \rangle$. But from the relation $S - xyzy''z = M + \langle xy'z \rangle + (N - \langle xy''z \rangle)$ it is clear that $(N - \langle xy''z \rangle)$ cannot be connected: for $\langle xy'z \rangle$ contains limit points both of M and of N, and if $N - \langle xy''z \rangle$ is connected, $\langle xy'z \rangle$ has a limit point of it also and the right side of the above relation is a connected point set. Then if the Jordan Curve Theorem is to fail to be true in S, it follows that x and z cannot both be accessible from N.

4. 3. Suppose then that one of these points, say x, is accessible from N. Then the other is not. But then it can be shown, as in § 3, that z is a point of a continu of condensation T such that every point of T is a limit point of N. Since T is a subset of K, and moreover, since $\langle xyz \rangle \cdot N = 0$, T is a subarc of $xy'z$: T is an arc tz. Then if t' is any point whatever of $\langle tz \rangle$, there is a point t'' of $\langle t'z \rangle$ and an arc xt'' such that $\langle xt'' \rangle \subset N$.

We may assume that the point t'' separates y' and z on $xy'z$. From $S - xyt''x = M + \langle xy't'' \rangle + (N - \langle xt'' \rangle)$ it is seen that $(N - \langle xt'' \rangle)$ cannot be connected, and has a component which fails to have any point of $\langle xy't'' \rangle$ for limit point. It is conceivable, however, that this domain fails, also, to have any limit point on $t''z$ ... xt''. We shall show that [illegible]

[illegible]

Let [illegible] such that $\langle xy \rangle$ [illegible]

[illegible]

$\langle x_1z_1\rangle \subset D_1$. Then $N - \langle xx_1z_1\rangle$ has a component $D' \subset D_1$ such that $\bar{D}' \cdot \langle xp_1z_1\rangle = 0$, but $\bar{D}' \cdot \langle z_1z\rangle \neq 0$. This contradicts our choice of z_1 unless $z_1 = z$; and this is not possible since z is supposed not accessible. Therefore z_1 is not accessible, and is seen to be a limit point of $\bar{D}_1 \cdot \langle p_1z_1\rangle$. We shall choose an arc x_2p_2, $x_2 \subset xp_1$ and $p_2 \subset p_1z$, and $\langle x_2p_2\rangle \subset D_1$, subject to the following conditions: (1) p_2 is a point of $\bar{D}_1 \cdot \langle p_1z_1\rangle$ such that the diameter of the arc p_2z_1 is less than $\epsilon/2$, where ϵ is the diameter of the arc p_1z_1; (2) if for some choice of p_2 subject to (1) and for some choice of x_2 on $\bar{D}_1 \cdot \langle xp_1\rangle$ (it is seen that this cannot be a vacuous set, since N is connected) there is an arc x_2p_2, $\langle x_2p_2\rangle \subset D_1$, whose diameter is d_2 then that arc x_2p_2 which we choose shall have a diameter not exceeding $d_2 + \epsilon/2$. We may express (2) in this form, that we choose a particular arc from among the set (non-vacuous) of available arcs so that its diameter does not exceed the lower limit of the diameters of the available arcs by more than $\frac{1}{2}\epsilon$: otherwise the choice of our arc is free (except as subject to (1)).

Then $N - \langle xx_2p_2\rangle$ has a domain $D_2 \subset D_1$ such that $\bar{D}_2 \cdot \langle xp_1p_2\rangle = 0$, but $\bar{D}_2 \cdot \langle p_2z\rangle \neq 0$. There is a point $z_2 \subset \bar{D}_2 \cdot p_2z \subset p_2z_2$. Let x_3p_3 be an arc, $x_3 \subset xp_2$, $p_3 \subset p_2z_2$, $x_3p_3 \subset D_2$, such that: (1) the diameter of the arc p_3z_2 is less than $\epsilon/4$; (2) the diameter of the arc x_3p_3 is at most $d_3 + \epsilon/4$, where d_3 is the lower limit of the diameters of the arcs (x_3p_3) which satisfy (1). We construct, inductively, for every n, an arc xx_np_n, a domain D_n, and a point z_n. From the fact that z_{m+1} either precedes z_m on $xy'z$, or is identical with it, it follows that the set of points (z_n) has a sequential limit z' which either precedes every point of (z_n) or is identical with all but a finite number of them. It follows from (1) that z' is also the sequential limit point of the set of points (p_n).

Suppose that infinitely many of the arcs (p_nx_n) are of diameter greater than some ϵ'. Since S is locally compact, we may suppose ϵ' such that $U_{z'\epsilon'}$ is contained in a compact subset of S. Then if $(p'_nx'_n)$ is the set of arcs of (p_nx_n) each of diameter greater than ϵ', where we suppose this subsequence such that for every n, $p'_n \subset U_{z'\epsilon'}$, there is on each arc in order from p'_n a first point q'_n which is not contained in $U_{z'\epsilon'}$. The set of arcs $(p'_nq'_n)$ has a subsequence $(p''_nq''_n)$ with a sequential continuum of condensation T. Since all but a finite number of the arcs of $(p''_nq''_n)$ are contained in $\bar{D}_m$ for every m, it is seen that no point of $\langle xp_{m-1}p_m\rangle + \langle z_mz\rangle$ can be a limit point of T. If then x is not contained in $U_{z'\epsilon'}$, T has no point in common with $xy'z$ excepting z'. Let t be a point of T distinct from z'. Then $t \subset \bar{D}_m$, for every m: we choose t moreover so that it is contained in $U_{z'\epsilon'/2}$. There is a k such that p_jx_j, $j > k$, is of diameter not

exceeding $d_j + \epsilon'/4$, where d_j is the lower limit of the diameters of the "available arcs $(p_j x_j)$".

Case 1: suppose $t \subset p''_i q''_i$, $i > k$. There is a neighborhood U_t of t such that any point of U_t may be joined to t by an arc of $U_{\epsilon'/2}$. Let t' be any point of an arc $p''_s q''_s$, $s > i$, such that $t' \subset U_t$. There is an arc $t't$ in $U_{1/2\epsilon'}$: this has a last point t'' on the arc $\langle p''_s q''_s \rangle$, and a first point t^* thereafter on $\langle p''_i q''_i \rangle$. Since $p''_s q''_s - p''_s \subset D''_i$, where D''_i is that one of the domains D_n which corresponds to the renumbered arc $p''_i q''_i$, it follows that $p''_s t'' t^*$ is contained, except for its endpoints, in D''_i. Say the domain D''_i is the domain D_r, that is, in the set (D_n). Then it is seen that the arc $p''_s t'' t^*$ corresponds to an arc $x_{r+1} p_{r+1}$ of diameter less than $\epsilon'/2$: therefore the arc $x_{r+1} p_{r+1}$ which was actually chosen should not have been of diameter greater than $3\epsilon'/4$. In this case we may conclude that not more than a finite number of the arcs $x_n p_n$ are of diameter greater than a preassigned ϵ.

Case 2: t is not a point of any arc xp_k. Then t is contained in D_n for every n. Let U_t be a neighborhood of t as above. Again, if $p''_s q''_s$ and $p''_i q''_i$ are arcs of $(p''_n q''_n)$ which have points t' and s' respectively in U_t, and $i > s > k$ (as above), we can obtain the contradiction of Case 1.

Then we may conclude that the arcs $x_n p_n$ have been so constructed that not more than a finite number of them are of diameter greater than any preassigned ϵ. From this, and the construction of the arcs xp_n, it is readily seen that the sum of all of the arcs (xp_n) is an acyclic continuous curve V whose points on $xy'z$ consist of x, z', and the points of (p_n). Then there is an arc xz' in this curve V which has no point other than x and z' in common with $xy'z$, because the points of (p_n) are endpoints of V. This arc xz', moreover, is contained in $\bar{D}_n$ for every n. There is a component of $X - \langle xz' \rangle$ which has no limit point in $\langle xy'z' \rangle$ and a limit point z'' in $\langle z'z \rangle$. It is seen that z'' is a limit point of every D_n. But for some n, points of z_n must precede z''. Since this is impossible because the point z_n is the last limit point which D_n has on $xy'z$, it follows that $z'' = z' = z$, and z is accessible. Therefore, our assertion is proved, that if x is accessible and z is not, there is for every point y' of $xy'z$, an arc xy_1 such that $y_1 \subset \langle y'z \rangle$, $xy_1 \subset X$, and that component of $X - \langle \cdot \rangle$ which has no limit point on $\langle xy'z \rangle$ [illegible]

[illegible]

p' and q' are two points of $xy'z$, there exists an arc pq, such that $pq \subset N$, and p and q lie on $xy'z$ in the order $xpp'y'qq'z$, and $S - pq$ is not connected: there is a domain $D_{pq} \subset N$.

4.4. Moreover if x, say, is not accessible then x cannot be a limit point of points of $\bar{M} \cdot xy'z$. For since x is not accessible there is an arc pq, where p may be the point z, such that $\langle pq \rangle \subset N$, $p + q \subset xy'z - z$, and $S - \langle pq \rangle$ contains a domain D which is a component of $N - \langle pq \rangle$. Let m' be a point of $\bar{M} \cdot \langle qx \rangle$, and let m be a point of $\bar{N} \cdot \langle m'x \rangle$ (such points exist, since x is not accessible). Then m is a limit point of a component of $N - \langle pq \rangle$ which has also a limit point m'' on the arc $\langle py'q \rangle$ of $xy'z$: i. e. there is only one domain of $N - \langle pq \rangle$ which fails to have a limit point on $py'q$—and in this case we have assumed that this domain has no limit point on $\langle qx \rangle$. Then there is an arc mm'' (we may suppose these points so chosen that they are accessible) such that $\langle mm'' \rangle \subset N$. If now mm'' has a domain D', i. e. if there is more than one component in $S - \langle mm'' \rangle$, or differently expressed, if $N - \langle mm'' \rangle$ has a component D' which, failing to have any limit point on the subarc $\langle m''m \rangle$ of $xy'z$, has no limit point also on either of the arcs $\langle zm'' \rangle$ or $\langle mx \rangle$, then the simple closed curve $pqmm''p$, where $\langle pq \rangle$ is the arc of N and mm'' as well as qm and $m''p$ are subarcs of $xy'z$, has at least three components. Then that domain D' of $N - \langle mm'' \rangle$ which has no limit point in $\langle m''m \rangle$ of $xy'z$ (such a domain exists from the existence of the point m' on $\langle mm'' \rangle$) must have a limit point on $\langle zm'' \rangle$ or $\langle xm \rangle$, and arguing within this domain we can show that there is no point which is either a last accessible point, or limit point z' of last accessible points, so that the accessibility of x must result.

4.5. We shall construct in M a simple closed curve which fails to separate S: the construction will resemble in numerous details the procedure of § 4.1.

Since every point of $\langle xy \rangle$ as well as of $\langle yz \rangle$ on K (see § 4) is a limit point of M, it is readily seen that there exists an arc hk, $\langle hk \rangle \subset M$, $h \subset \langle xy \rangle$, $k \subset \langle yz \rangle$. Then $S - hxy'zkh = N + \langle hyk \rangle + (M - \langle hk \rangle)$, and it follows that $\langle hyk \rangle + (M - \langle hk \rangle)$ is connected: in fact it is an $(hxy'zkh)$-domain. Since y' is a limit point of M, it is a limit point of $(M - \langle hk \rangle)$, therefore of $\langle hyk \rangle + (M - \langle hk \rangle)$. There exists an arc st, $s \subset \langle hyk \rangle$, $t \subset \langle xy'z \rangle$, $st \subset (M - \langle hk \rangle)$: $(st) \cdot (hk) = 0$. Since $S - stxs = N + \langle szt \rangle + (M - \langle st \rangle)$, there is a component A of $(M - st)$ which has no limit point in $\langle szt \rangle$, and A is the only such component: for if no such component exists, $S - stxs$ is connected, and if more than one then there are in $S - stxs$

at least three domains. If $\bar{A} \subset A + st$, let pq be an arc such that $\langle pq \rangle \subset N$, $p \subset \langle xt \rangle$, $q \subset \langle tz \rangle$. If, now, $N - \langle pq \rangle$ has any component D such that $\bar{D} \subset D + pq$, the simple closed curve $stpqzs$, where tp and qzs are arcs of K, has at least three domains. Then no such arc pq exists, and either x or z must be accessible. If now x, say, is accessible let xq be an arc such that $\langle xq \rangle \subset N$, $q \subset \langle tz \rangle$, and $N - \langle xq \rangle$ has a domain D' such that $\bar{D}' \subset D' + xq$: then $stxqzs$ has three domains, and z also is accessible. But we have previously shown that x and z cannot both be accessible. Therefore it follows that A has at least one limit point in $\langle sxt \rangle$.

Since $S - stzs = N + \langle sxt \rangle + (M - \langle st \rangle)$, it is seen that $N - \langle st \rangle$ has a component B, and only one, such that $\bar{B} \cdot \langle sxt \rangle = 0$; it follows the argument above that $B \cdot \langle szt \rangle \neq 0$, and it is clear that B and A are distinct. Moreover, since $hk \cdot st = 0$ and therefore $\langle hk \rangle \subset (M - \langle st \rangle)$ and has a limit point on both $\langle sxt \rangle$ and $\langle szt \rangle$, it is seen that $M - \langle st \rangle$ has at least one other component C. We shall show, however, that $M - \langle st \rangle$ cannot have more than the two components A and B. Suppose that no point of $xs - s$ is a limit point of A. Then there is on $\langle xt \rangle$ a point x'' such that $x''x \cdot \bar{A} = 0$. Let pq be an arc, $\langle pq \rangle \subset N$, $p \subset x''x$, $q \subset tz - t$. Since $\bar{A} \subset A + \langle stx'' \rangle$, it is seen as in the foregoing paragraph that the component of $N - \langle pq \rangle$ which has no limit point in $\langle ptq \rangle$ cannot fail to have limit points in $xp - p$ or in $zq - q$, and that in consequence both x and z must be accessible from N (for in the arc pq, p may be x or q may be z). Then $(xs - s) \cdot \bar{A} \neq 0$: suppose, however, that $\langle xs \rangle \cdot \bar{A} = 0$. If x is not a limit point of points of $\bar{A} \cdot \langle tx \rangle$, it readily follows that x must be accessible from A. If x is a limit point of points of $\bar{A} \cdot \langle tx \rangle$ it follows that x is accessible from N (§ 4. 4) and then that z is accessible, since p may be x in the arc, above, pq. Therefore either x is accessible from A, or there is at least one point in $\bar{A} \cdot \langle sx \rangle$, and consequently a point $x' \subset \bar{A} \cdot \langle sx \rangle$ such that x' *is* accessible from A. It will be immaterial to the sequel whether the point x' or x is accessible from A, and therefore we shall write x' with the understanding that x' is a point of $xs - s$, and may be x.

By an entirely analogous argument there is a point z' on $zs - s$ such that z' is accessible from B. Let a be any point of $\bar{A} \cdot \langle st \rangle$ such that there is an arc ax', $\langle ax' \rangle \subset A$: the existence of such a point readily follows from this, that $\bar{A} \cdot \langle st \rangle$ cannot be vacuous. Suppose that $\bar{B} \cdot \langle at \rangle = 0$: the simple closed curve $ax'tz'sa$ has at least three domains, one of them N, the other B, and a third containing $\langle sx' \rangle$, for example. Then there is a point b of $\bar{B} \cdot \langle at \rangle$ such that there is an arc bz', $\langle bz' \rangle \subset B$. Let H be that $(bz'tx'sab)$-domain which does not contain N. Since $S - bz'tx'sab = N + \langle sz' \rangle + (M - \langle sbz' \rangle)$, H must contain the set $\langle sz' \rangle + (M - \langle sbz' \rangle)$. There is

then an arc fg of H, $fg \cdot \langle bt \rangle = f$ and $fg \cdot \langle sz' \rangle = g$. Then $\langle fg \rangle$ belongs to a component of $M - \langle st \rangle$, and this component cannot be A.

Case 1: Suppose $\langle fg \rangle \subseteq B$. If $\bar{A}$ has no limit point in $\langle ft \rangle$, the simple closed curve $fgz'tx'sf$ has at least three components: then there is an arc $a'ka$, $\langle a'ka \rangle \subset A$, $a' \subset \langle ft \rangle$, and $a'k \cdot ax' = k$. Let H' be that $btz't$-domain which does not contain N. Then H' does not contain the arc $txsz'$, or the arc sb, or the arc $z'f$: $H' \subset M - \langle st \rangle$ and belongs to a component of $M - \langle st \rangle$ which has no limit point in $\langle sxt \rangle$; and $H' \subset B - \langle bz' \rangle$. Since g is a limit point of B, therefore of $B - \langle bz' \rangle$, it follows that H' is a proper subset of $B - \langle bz' \rangle$ and has at least one limit point in $\langle bz' \rangle$. Moreover, if $H' \cdot \langle a't \rangle = 0$, the simple closed curve $a'kx'sz'ba'$ has at least three components, H', N, and a third which contains, for example, $\langle x'sz' \rangle$. Then there is an arc $b'h$, $\langle b'h \rangle \subset H'$, $b' \subset \langle a't \rangle$, and $h \subset \langle bz' \rangle$. Let H'' be that $aka'b'hba$-domain which does not contain N. Then H'' contains no point of $sxtzs + sa + gf + bfa' + b't$: being a subset of $M - \langle st \rangle$ it belongs to a component of $M - \langle st \rangle$. Since every component of $M - \langle st \rangle$ other than A and B is maximally connected in $S - aka'b'hba$ and has at least one limit point in $sxtzs$, no component other than A and B can have any point in H''. Then $H'' \subset A + B$, and since $H'' \cdot \langle st \rangle = 0$, H'' belongs to A or H'' belongs to B. In the first case the simple closed curve $baka'b'hz'tx'sgfb$ has at least three components, and in the second the simple closed curve $bhb'a'kx'tz'gfb$.

Case 2: $\langle fg \rangle$ does not belong to B. Then $\langle fg \rangle$ is contained in a component C' of $M - \langle st \rangle$, distinct from B and from A. As in Case 1, there is an arc $a'ka$, $\langle a'ka \rangle \subset A$, $a' \subset \langle ft \rangle$, $a'k \cdot x'a = k$. If B has no limit point in $\langle a't \rangle$, the simple closed curve $a'kx'tz'sa'$ has at least three domains. Since $\langle bz' \rangle \subset B$, there is an arc $b'h$, $\langle b'h \rangle \subset B$, $b' \subset \langle a't \rangle$, $h \subset \langle bz' \rangle$. The simple closed curve $aka'b'hba$ has no point in common with $\langle fg \rangle$ since this belongs to C', and we arrive at a contradiction as in Case 1.

Then we have shown that S satisfies the Jordan Curve Theorem, and S is a cylinder tree. We have the

THEOREM 4. *A necessary and sufficient condition that a continuous curve be a cylinder tree is that it be cyclicly connected and satisfy the property* (G): *if K is a simple closed curve of S, then every point of K is a limit point of $S - K$ and $S - K$ is the sum of precisely two components.*

5. We shall show that the Theorem 4 remains true if the condition of

cyclic connectivity be replaced by the weaker condition that S contains at least one simple closed curve (property A).

For if S contains a simple closed curve, it contains a maximal cyclic curve J. If any two points x and y of J are cut-points of S, and K is a simple closed curve of J containing x and y, it is seen that $S - K$ has at least three components. Therefore J has at most a single cut-point x of S. On our definition of a continuous curve, $J - x$ is a continuous curve J'. If any point y of J' is a cut-point of J', and if K is a simple closed curve of J containing x and y, it is readily seen that $S - K$ contains at least three distinct components. Therefore J' is cyclicly connected. If K' is any simple closed curve of J', every point of K' being a limit point of $S - K'$, but not a limit point of $S - J'$, is necessarily a limit point of $J' - K'$; and, moreover, if $J' - K'$ consists of more than two components it follows that $S' - K'$ consists of more than two components. Then J' has the properties (D) and (G) and by the theorem above is a cylinder tree.

Let S' be the simple closed surface such that J' is homeomorphic with $S' - B$, where B is a closed and totally disconnected point set. Since x of J is a limit point of J', there is a sequence of points of J' converging to x, and the corresponding points on S' converge to a point x' of S' which is necessarily a point x' of B. Moreover since J is locally compact, in a neighborhood of x every infinite set of points of J' has a limit point—which may be x. Then in a neighborhood of x' on S', every infinite set of points of $S' - B$ has a limit point which may be x', but is not a point of $B - x'$. If the homeomorphism be extended to x and x', that is if x and x' be defined as corresponding points in this transformation, it follows that there exists a homeomorphism between J and $S' - (B - x')$, and $B - x'$ is totally disconnected and closed. That is, J is a cylinder-tree. If then, K'' is a simple closed curve of $S' - (B - x')$ which contains x', it separates $S' - (B - x')$ into two distinct components, and if K^* is the corresponding simple closed curve of J, containing x, $J - K^*$ is the sum of two distinct components. Then it follows that x cannot be a cut-point of S, or $S - K^*$ contains at least three components. Then S is also cyclicly connected, and is a cylinder tree.

THEOREM 5. *A necessary and sufficient condition that a continuous curve be a cylinder-tree is that it contain at least one simple closed curve and have the property G.*

6. Suppose now that S is any space satisfying the following Axioms: Axiom 1, 4, and 5 of $\Sigma 1$.*

* F. A.

Axiom $2'$: a region is a connected set of points, not a single point.

Axiom $3'$: if x is a point of S, $S - x$ is connected.

Axiom $8'$: If K is a simple closed curve of S, every point of K is a limit point of $S - K$, and $S - K$ is the sum of two components at least one of which is compact.

From the work of E. W. Chittenden † it follows that S, satisfying Axioms 1 and $2'$ is a metric separable space. From Axiom $3'$ it follows that S is connected, and from Axiom 4 that S is locally compact. From Axioms 1 and $2'$ it follows, readily, that S is connected im kleinen; with Axiom $3'$ that S is cyclicly connected. Axiom 5 asserts that S is not compact, Axiom $8'$ that every simple closed curve of S determines a compact domain, and that S has moreover the property G. Then S is a cylinder tree, and moreover a number-plane.

It is possible, in view of Theorem 5, to replace Axiom $3'$ by either of the following Axioms:

Axiom $3''$ S contains at least one simple closed curve.

Axiom 3^* if m is any arc of S, then at least one point of m fails to disconnect S.

We have, finally, the theorem:

THEOREM 6. *A necessary and sufficient condition that a space S be a number-plane is that it satisfy the Axioms,* 1, $2'$, $3'$, 4, 5, *and* $8'$.

AUSTIN, TEXAS.

† "On the Metrization Problem and Related Problems in the Theory of Abstract Sets," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), p. 22, No. 8.

The Linear Element of a Riemannian V_n in Terms of the Christoffel Symbols of the Second Kind.*

By W. C. Graustein.

1. *Introduction.* The problem of determining conditions necessary and sufficient that there exist a nonsingular quadratic differential form in n variables $x^1, x^2, \cdots, x^n$:

$$ds^2 = g_{ij}dx^i dx^j, \qquad g_{ij} = g_{ji}, \qquad g \equiv | g_{ij} | \neq 0, \tag{1}$$

which has as its Christoffel symbols $\left\{ \begin{matrix} i \\ jk \end{matrix} \right\}$ prescribed functions, $\Gamma^i{}_{jk}$, has been successfully treated in all its generality.† The solution obtained, though ostensibly the best possible for the unrestricted problem, is involved and fails to furnish conditions which can readily be expressed explicitly in terms of the given functions $\Gamma^i{}_{jk}$.

It is the purpose of this paper to show that if $n = 2$, or if (1) is restricted to be the linear element of a Riemannian space of constant Riemannian curvature or, more generally, of an Einstein space, conditions of a very simple nature, expressed directly in terms of the given functions $\Gamma^i{}_{jk}$, can be found.

The general analytic problem has various geometric interpretations. The conditions sought are the conditions necessary and sufficient (a) that the functions $\Gamma^i{}_{jk}$ be the Christoffel symbols of the second kind of a Riemannian space V_n; (b) that the differential equations

$$\frac{d^2x^i}{ds^2} + \Gamma^i{}_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0, \qquad (i = 1, 2, \cdots, n)$$

be those of the geodesics of a Riemannian space V_n and s be the arc of an arbitrary geodesic; (c) that the space V_n with the coefficients of connection $\Gamma^i{}_{jk}$ be a Riemannian space.

We shall find it convenient to formulate our treatment and results in terms of the geometric language of the last of these interpretations.

2. *Preliminary Considerations.* Since $\left\{ \begin{matrix} i \\ jk \end{matrix} \right\}$ is symmetric in j and k, necessary conditions on the functions $\Gamma^i{}_{jk}$ are

$$\Gamma^i{}_{jk} = \Gamma^i{}_{kj}.$$

* Presented to the Society, December 28, 1926.

† Cf. Eisenhart, "Non-Riemannian Geometry," American Mathematical Society Colloquium Publications, Vol. VIII, § 29.

We assume henceforth that these conditions are always satisfied, i. e. that the connection Γ^i_{jk} is always symmetric.

The curvature tensor of the space with the symmetric connection Γ^i_{jk}, namely

$$B^i_{jkl} = (\partial\Gamma^i_{jl}/\partial x^k) - (\partial\Gamma^i_{jk}/\partial x^l) + \Gamma^i_{hk}\Gamma^h_{jl} - \Gamma^i_{hl}\Gamma^h_{jk},$$

and its contraction, $B_{jk} = B^i_{jki}$, are the generalizations respectively of the Riemannian curvature tensor R^i_{jkl} and the Ricci tensor R_{jk} of a Riemannian space. Since the tensor R_{ij} is symmetric, necessary conditions on Γ^i_{jk} are that the tensor B_{ij} be symmetric.

When $B_{ij} = B_{ji}$, the projective curvature tensor of Weyl has the form

$$W^i_{jkl} = B^i_{jkl} - [1/(n-1)]\ (\delta_l{}^i B_{jk} - \delta_k{}^i B_{jl}),$$

and the relations

$$B_{ih}W^h_{jkl} + B_{hj}W^h_{ikl} = B_{ih}B^h_{jkl} + B_{hj}B^h_{ikl}, \tag{2}$$

which will prove of use later, are readily established.

3. *A General Theorem.* As the point of departure for the consideration of the general problem, we employ the well known proposition.*

LEMMA 1. *A necessary and sufficient condition that the functions* Γ^i_{jk}, *symmetric in* j *and* k, *be the Christoffel symbols of the second kind of the differential form* (1) *is that*

$$g_{ij,k} = 0, \tag{3}$$

where the covariant differentiation is with respect to Γ^i_{jk}.

If for g_{ij} we set

$$g_{ij} = e^{-\phi}a_{ij},$$

where $\phi(x^1, x^2, \cdots, x^n)$ is an invariant function, equations (3) become

$$a_{ij,k} = \phi_{,k}a_{ij}, \tag{4}$$

where $\phi_{,k}$ is the gradient of the function ϕ.

If a tensor a_{ij} satisfies equations of the form (4), it obviously satisfies equations of the form

$$a_{ij,k} = \phi_k a_{ij}, \tag{5}$$

where ϕ_k is a covariant vector. It also satisfies the equations

$$a_{ih}B^h_{jkl} + a_{hj}B^h_{ikl} = 0, \tag{6}$$

resulting from the conditions of compatibility,

$$a_{ij,kl} - a_{ij,lk} = a_{ih}B^h_{jkl} + a_{hj}B^h_{ikl},$$

of the equations (4).

---

* Cf. Eisenhart, *loc. cit.*, § 29.

Conversely, if a tensor a_{ij} satisfies equations of the form (5) and also equations (6), it satisfies equations of the form (4). For, if a_{ij} satisfies (5), it satisfies the equations of compatibility of (5), which reduce to

$$a_{ij}(\phi_{k,l} - \phi_{l,k}) = 0.$$

Hence, if $a_{ij} \not\equiv 0$, ϕ_k is the gradient of a function $\phi : \phi_k = \phi_{,k}$, and a_{ij} satisfies (4). On the other hand, a tensor a_{ij} all of whose components are zero surely satisfies (4).

Since equations (5) and (6) are equivalent to equations (4) and hence to equations (3), we conclude:

THEOREM 1. *A necessary and sufficient condition that there exist a tensor $g_{ij} \not\equiv 0$ satisfying equations* (3) *is that there exist a tensor $a_{ij} \not\equiv 0$ which satisfies the equations*

$$a_{ij,k} = \phi_k a_{ij}, \qquad a_{ih}B^h{}_{jkl} + a_{hj}B^h{}_{ikl} = 0, \tag{7}$$

where ϕ_k is a covariant vector. Then ϕ_k is the gradient of a function ϕ, and $g_{ij} = ce^{-\phi}a_{ij}$, where c is an arbitrary constant, satisfies (3).

We have also:

LEMMA 2. *A necessary and sufficient condition that equations* (3) *possess a solution of the form*

$$g_{ij} = \rho a_{ij},$$

where a_{ij} is a given tensor, $\not\equiv 0$, and ρ an invariant function, $\not\equiv 0$, is that a_{ij} satisfy equations (7). *Then ϕ_k is the gradient of a function ϕ and*

$$g_{ij} = ce^{-\phi}a_{ij}, \qquad c \neq 0,$$

is the general solution of (3) *of the prescribed form.*

The first set of equations (7), i. e. equations (5), may be replaced by the equivalent equations

$$a_{rs}a_{ij,k} - a_{ij}a_{rs,k} = 0.$$

In other words, equations (5) are to be thought of as demanding merely that $a_{ij,k}$ be equal to the product of a_{ij} and *some* covariant vector ϕ_k—not one which is prescribed. The same significance is to be ascribed to [illegible]

[illegible]

of a Riemannian space V_2 is a multiple of the Ricci tensor R_{ij}, provided that the curvature K of the space does not vanish identically. In fact,

$$g_{ij} = -(1/K)R_{ij}, \qquad K \neq 0. \tag{8}$$

Hence, a space V_2 with the symmetric connection Γ^i_{jk} is a Riemannian space of non-zero curvature if and only if $B_{ij} = B_{ji}$, $|B_{ij}| \neq 0$, and equations (3) have a solution of the form $g_{ij} = \rho B_{ij}$, $\rho \neq 0$. By Lemma 2, the last of the conditions is equivalent to

$$B_{ij,k} = \phi_k B_{ij}, \qquad B_{ih}B^h{}_{jkl} + B_{hj}B^h{}_{ikl} = 0.$$

But the equations of the second of these sets are identically satisfied, by virtue of the relations (2) and the fact that the Weyl tensor $W^i{}_{jkl}$ vanishes identically when $n = 2$.

THEOREM 2. *A necessary and sufficient condition that a space V_2 with symmetric connection Γ^i_{jk} be a Riemannian space of non-zero curvature is that*

$$B_{ij,k} = \phi_k B_{ij}, \qquad B_{ij} = B_{ji}, \qquad |B_{ij}| \neq 0.$$

The fundamental tensor g_{ij} can then be found by quadratures and is uniquely determined to within a constant factor.

By Lemma 2 and equations (8), we have

$$g_{ij} = ce^{-\phi}B_{ij}, \qquad K = -(1/c)e^{\phi}, \qquad c \neq 0.$$

Hence, $B_{ij,k} = 0$ if and only if $\phi_k = 0$ or $K = \text{const.}$

COROLLARY. *A necessary and sufficient condition that the given space be a Riemannian space of constant curvature, not zero, is that*

$$B_{ij,k} = 0, \qquad B_{ij} = B_{ji}, \qquad |B_{ij}| \neq 0.$$

5. *Einstein Spaces.* A Riemannian space V_n for which the Ricci tensor R_{ij} is a multiple ρ of the fundamental tensor g_{ij} is known as an Einstein space. The multiplier ρ is necessarily equal to R/n, where $R = g^{ij}R_{ij}$ is the scalar curvature of the space. Thus, an Einstein space of non-zero scalar curvature is completely characterized by the relations

$$g_{ij} = (n/R)R_{ij}, \qquad R \neq 0. \tag{9}$$

By an argument similar to that employed in the case $n = 2$, we arrive at the following result.

THEOREM 3. *A necessary and sufficient condition that a space V_n with the symmetric connection Γ^i_{jk} be an Einstein space of non-zero scalar curvature is that $B_{ij} = B_{ji}$, $|B_{ij}| \neq 0$, and*

$$B_{ij,k} = \phi_k B_{ij}, \qquad B_{ih} W^h{}_{jkl} + B_{hj} W^h{}_{ikl} = 0.$$

The fundamental tensor g_{ij} can then be found by quadratures and is uniquely determined to within a constant factor.

We have, in fact,

$$g_{ij} = ce^{-\phi} B_{ij}, \qquad R = (n/c)\, e^{\phi}, \qquad c \neq 0.$$

When $B_{ij,k} = \phi_k B_{ij}$ is replaced by $B_{ij,k} = 0$, the conditions become those under which the given space is an Einstein space of constant scalar curvature, not zero. The corollary thus obtained is, however, of importance only in the case $n = 2$, since every Einstein space for $n > 2$ is of constant scalar curvature.

6. *Riemannian Spaces of Constant Curvature.* When $n > 2$, the Riemannian curvature of a Riemannian space V_n is constant if and only if the Weyl tensor, formed for the moment in terms of $R^i{}_{jkl}$ and R_{ij}, vanishes identically. But when $n = 2$, $W^i{}_{jkl}$ is always identically zero. Consequently, a necessary and sufficient condition that the Riemannian curvature of a Riemannian space V_n be a point function (and hence, when $n > 2$, a constant) is that $W^i{}_{jkl} = 0$.

The Riemannian spaces V_2 and the Riemannian spaces V_n, $n > 2$, of constant Riemannian curvature are Einstein spaces. We have, in fact, in both cases,

$$g_{ij} = -[1/(n-1)K]\, R_{ij}, \tag{10}$$

provided the Riemannian curvature K is not zero.

In light of these considerations, we conclude immediately, from Theorem 3:

THEOREM 4. *A necessary and sufficient condition that a space V_n with the symmetric connection $\Gamma^i{}_{jk}$ be a Riemannian space whose Riemannian curvature is a point function, not zero, is that $B_{ij} = B_{ji}$, $|B_{ij}| \neq 0$, and*

$$B_{ij,k} = \phi_k B_{ij}, \qquad W^i{}_{jkl} = 0.$$

In this case, we have

$$g_{ij} = ce^{-\phi} B_{ij}, \qquad K = -[1/(n-1)c]\, e^{\phi}, \qquad c \neq 0.$$

Here, too, there is a corollary involving the replacement of the condition $B_{ij,k} = \phi_k B_{ij}$ by $B_{ij,k} = 0$, but, as in the previous case, the corollary has significance only when $n = 2$.

It is to be noted that Theorem 4 is a special case of Theorem 3 and Theorem 2 in turn a special case of Theorem 4.*

7. *Projectively Flat Spaces.* The space V_n with the symmetric connection Γ^i_{jk} is called (absolutely) flat if its curvature tensor B^i_{jkl} vanishes, i. e. if it is identical with a Riemannian space of zero Riemannian curvature. For reasons which we need not go into here, the space V_n is said to be projectively flat if it can be put into one-to-one point correspondence with a flat space so that to the paths of the one space correspond the paths of the other.

The condition $W^i_{jkl} = 0$ of Theorem 4 is necessary and sufficient that the given space V_n be projectively flat when $n > 2$, and, when, $n = 2$, it imposes no restriction.

Furthermore, if a projectively flat space V_n is Riemannian, it can be put into one-to-one point correspondence with a Riemannian space of constant (zero) Riemannian curvature so that geodesics correspond, and hence is itself of constant Riemannian curvature, by the theorem of Beltrami.

THEOREM 5. *A necessary and sufficient condition that a projectively flat space V_n with the symmetric connection Γ^i_{jk} be a Riemannian space of non-zero Riemannian curvature is that*

$$B_{ij,k} = \phi_k B_{ij}, \qquad B_{ij} = B_{ji}, \qquad | B_{ij} | \neq 0.$$

If the conditions $B_{ij,k} = \phi_k B_{ij}$ are replaced by $B_{ij,k} = 0$, the theorem remains valid, for in this case the Riemannian space must be of constant Riemannian curvature for $n = 2$ as well as for $n > 2$. The alternative conditions are, however, more restrictive than the original ones.†

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

* The author is indebted to J. M. Thomas for pointing out that his original method of treatment, which was confined to the case $n = 2$, could be applied equally well, when $n > 2$, to projectively flat spaces. The present method of treatment, based on the general Theorem 1 and the relations (2), resulted from generalizing the original method so that it would apply also to the case in which (1) is assumed to be the linear element of an Einstein space.

† The alternative theorem is known; cf. Eisenhart, *loc. cit.*

The Regular Components of Surface Transformations.

By Paul A. Smith.

On a closed surface which is subjected to a one-one continuous transformation into itself, there exists a certain closed invariant point set towards which the remaining points of the surface are carried when the transformation is indefinitely iterated. In a variety of cases, it is possible to determine on the surface certain open sets characterized by a property of "uniform motion" towards the invariant central set. These so-called regular components were first considered and their structure studied by Birkhoff and Smith,* but only for the case in which the surface was topologically equivalent to a sphere; in this case the treatment is relatively simple since it is possible to avoid extensive reference to the facts and methods of surface topology. In the present study of regular components, we shall make virtually no restriction on the surface undergoing transformation save that it be closed. Regular components will be classified with regard to their connectivity and their motion under iteration. It will appear that the number of possible types of components is independent of the connectivity of the surface on which they exist. With one exception, each type of component found on a given surface may equally well exist on every other surface. Transformations which admit regular components of the exceptional type are of very special structure; they are analyzed in some detail in the final section.

1. *Notation and definitions.* We shall consistently denote by S a closed surface of arbitrary connectivity and orientability, and by T an arbitrary $(1, 1)$ continuous transformation of S into itself.

In introducing certain definitions we shall require the use of a metric on S. For definiteness we shall suppose S to be a suitably regular surface situated in a Euclidean space R_n and to derive its metric from that of R_n. The distance between the two points of S will be taken to mean the geodesic distance on S; there will be no ambiguity since only arbitrarily small distances

subset of S which constitutes a simple closed curve in R_n, and a region on S will mean a connected subset of S, the points of which are inner points relative to S.

A region on S which is homeomorphic to a region in a Euclidean plane, is a region of *planar type.* In particular, if a region R is homeomorphic to the plane region interior to a circle, R is simply connected, and if R is homeomorphic to the region between two concentric circles, R is *doubly connected.* A doubly connected region on S whose boundary consists of two non-intersecting circuits will be called a *ring.*

In recalling the following facts concerning the connectivity of regions on S, we are drawing freely on well established theorems of surface topology.*

A necessary and sufficient condition that a region R on S be of planar type is that R shall be a proper subset of S, and that an arbitrary circuit α in R shall separate it into two regions; and a necessary and sufficient condition that R be simply connected is that for every choice of α, one of the two regions shall be a 2-cell.

If R is doubly connected, there can of course be drawn in R a circuit which is not the boundary of a 2-cell contained in R. If R is of planar type, but not simply or doubly connected (*i. e.*, is of "higher order of connectivity"), there can be drawn in R a pair of non-intersecting circuits neither of which bounds a 2-cell in R, and which together fail to bound a ring in R.

2. We shall recall a few additional facts of surface topology preliminary to proving a useful lemma.

Let U be a surface either closed or with boundary.† A fundamental topological invariant of U is its characteristic $C(U) = \alpha_0 - \alpha_1 + \alpha_2$ where α_0, α_1, and α_2 are respectively the numbers of 0—, 1—, and 2-cells which form an arbitrary subdivision of U. If U possesses a boundary, then $C(U) \leq 1$. In particular if $C(U) = 1$, then U (minus its boundary) is equivalent either to a 2-cell or to a Möbius strip; and if $C(U) = 0$, U is a ring (§ 1). For all other types of surfaces with boundary, $C(U) < 0$.

Suppose now that S be separated by means of a finite number of non-intersecting 2-sided circuits into two or more regions $V^1, \cdots, V^n$. If we call U^i the surface with boundary consisting of V^i plus its boundary points, it is readily verified that

$$C(S) = C(U^1) + \cdots + C(U^n).$$

* See, for example, Kerékjártó, "Vorlesungen über Topologie I."

† A surface with boundary is homeomorphic to a closed surface from which there have been removed a finite number of 2-cells the boundaries of which are mutually exclusive.

Hence if k of the regions V^i are Möbius strips, the contribution of the corresponding U's toward $C(S)$ is k. Let us assume that none of the remaining $n-k$ regions is a 2-cell or a ring. Then the contribution of the corresponding U's is $\leq(n-k)$ and we have $C(S) \leq k-(n-k)$ and $n \leq 2k - C(S)$. Now the largest possible value k' of k relative to all possible subdivisions of the type considered depends only on the connectivity index of S. Hence, under the assumption that no region of the subdivision is a ring or a 2-cell, we have that $n \leq \eta(S) = 2k' - C(S)$; $\eta(S)$ depends only on the topological invariants of S. Consequently, if a given S be separated into more than $\eta(S)$ regions, one at least is a 2-cell or a ring.

From this we readily prove the following

LEMMA. *For a given S, there can be determined an integer $N > 1$ such that of every set of N or more mutually exclusive 2-sided non-cell-bounding circuits on S (if any such exist), at least one pair bounds a ring on S.*

Proof. It is a simple exercise in surface topology to show that given an integer $K > 1$, there can be chosen an integer L such that every set of L or more mutually exclusive circuits of the stated type on S separate S into at least K regions. If in particular $K = \eta(S) + 1$, one of the regions is a ring and the corresponding value for L may therefore be taken for N.

It is clear that the smallest possible value N^* of N for a given S is the same for every surface homeomorphic to S, and depends therefore only on the invariants of S. The explicit form of this dependence can easily be derived. It will be sufficient for our purposes, however, to point out that for a torus, $N^* = 2$; in fact, a torus is separated by every pair of non-intersecting non-cell-bounding circuits into two rings.

3. *Regular points.* We shall now state briefly the essential definitions and theorems relative to the motion of points under indefinite iteration. For more completeness of detail, the paper referred to above (S. A.) may be consulted.

Let $T_2, T_3, \cdots$ be the successive powers of $T(=T_1)$ and $T_{-2}, T_{-3}, \cdots$ those of the inverse T_{-1}. If E is an arbitrary point set on S, we shall denote by E_n $(n = \pm 1, \pm 2, \cdots)$ the transformed set $T_n(E)$.

Suppose that a given region σ on S has the property that it is intersected by none of its images $\cdots \sigma_{-2}, \sigma_{-1}, \sigma_1, \sigma_2, \cdots$. Then no two images of σ can intersect. For if σ_p and σ_q $(p \neq q)$ have points in common, so do σ and σ_{p-q}. Any region of type σ on S is called a *wandering region*, and any point contained in a wandering region is a *wandering point*; the totality of

wandering points constitutes an open set W. The closed set $M = S - W$ is non-null and consists of the *non-wandering* points of S; M is invariant under T,—that is, M and $T(M)$ are identical point sets (S. A., p. 351).

An important property of wandering points is the following: if P is an arbitrary wandering point and (P_n, M) is the distance from P_n to M, then $\lim_{n\to\pm\infty} (P_n, M) = 0$ (S. A., p. 351), or, as we shall say, P converges toward M under indefinite iteration of T as well as T_{-1}.

We shall be interested in the motion of sets of wandering points, especially as regards the uniformity of their convergence toward M. Suppose that a region σ of wandering points has the property that there exists a positive integer h such that the set $\sigma_h + \sigma_{h+1} + \cdots$ (or $\sigma_{-h} + \sigma_{-h-1} + \cdots$) is entirely contained in an arbitrarily small neighborhood V of M, h depending only on the choice of V. We shall call a point contained in such a region an ω- (*or* α-) *regular point;* wandering points contained in no such region are ω- (α-) *irregular.*

Suppose that a closed set E consist entirely of ω- (α-) regular points. It follows from an application of the Heine-Borel theorem that for $n \geqq N$ each point of $E_n(E_{-n})$ is contained in an arbitrarily small neighborhood V of M, where N depends only on the choice of V.

The totality of ω- (α-) regular points, if any exist, constitutes an open set, any maximal connected subset of which we shall call an ω- (α-) *regular component.** The image under T or T_{-1} of an ω- (α-) regular component is again an ω- (α-) regular component and from this it follows readily that a regular component C is either a wandering region, or else there exists an integer $k > 0$ such that C_p and C_q are identical if $p \equiv q \pmod{k}$ and mutually exclusive if $p \not\equiv q \pmod{k}$. We shall call regular components of the latter type *periodic,* of order k. It is our purpose to study the structure of components of both types.

THEOREM 1. *If the boundary ρ of a region R on S consists of ω-regular points, and if for some integer $h > 0$ the regions R, R_h, R_{2h}, $\cdots$ are mutually exclusive, then each point of R is ω-regular. A similar theorem holds for the α-regular case.*

Proof. The regions in each of the h sequences

$$R_i,\ R_{i+h},\ R_{i+2h},\quad (i = 0, 1, \cdots, h-1)$$

* We shall frequently use the term "regular" with reference to a point or component which is ω- or α-regular or both. In all proofs, however, "regular" will mean ω-regular for definiteness.

are mutually exclusive. Since the surface measure mS of S is finite (see § 1), it follows that as $n \to \infty$, lim $m R_{i+nh} = 0$ for $i = 0, 1, \cdots h - 1$, and hence lim $mR_n = 0$. Therefore, if ϵ is an arbitrary positive number, each point of R_n is within a distance $< \epsilon$ from ρ_n whenever n is greater than a suitably chosen positive integer N. Since the points of the closed set ρ are ω-regular, each point of ρ_n is within the distance ϵ from M, whenever n is greater than K, suitably chosen. Each point of R_n is therefore within a distance of 2ϵ from M whenever $n > N + K$. It follows that all wandering points contained in R are ω-regular.

The proof of the theorem will be complete if we show that each point of R is wandering. Suppose on the contrary that R contains points of M. Then (S. A., theorem 4, p. 360.) R is intersected by infinitely many of its images $R_1, R_2, \cdots$. Let one of these intersecting images be R_p, where p is chosen greater than $N + K$ so that each point of R_p is within a distance 2ϵ from M. Now ρ is at a non-zero distance δ from M, and ϵ being arbitrary, we may assume that $2\epsilon < \delta$. Hence R_p fails to meet ρ and is therefore entirely contained in R. We have then that $R \supseteq R_p \supseteq R_{2p} \supseteq \cdots \supseteq R_{hp}$. But this is impossible, since R and R_{hp} are mutually exclusive. This completes the proof.

THEOREM 2. *If a component D (i. e., a maximal connected subset) of $S - M$ is a wandering region, then D is an ω- and α-regular component.*

The proof is similar to that of the preceding theorem. We have $\lim_{n \to \pm\infty} mD_n = 0$. Hence, since the boundary of each D_n is in M, the totality of points in any region contained in D is brought within and remains within an arbitrarily small distance of M on indefinite iteration of T or T_{-1}.

4. *Generating Rings.* Suppose there exists on S a ring r whose boundary circuits α and β are transforms one of the other under a power of T, say $\beta = \alpha_m$ $(m > 0)$. Suppose further that r is such that the rings $r, r_m, r_{2m}, \cdots$ are mutually exclusive. We shall call r a *generating ring* of order m. It is clear that the rings of the sequence

[illegible]

[illegible]

THEOREM 3. *Neither circuit of the boundary of a generating ring can be entirely contained in a regular component C of wandering type.*

Suppose on the contrary that $\alpha + \alpha_m$ is the boundary of a generating ring r and that α is contained in C. Then since the points of α and α_m are regular, so are those of r by theorem 1, and r therefore contains no boundary points of C or C_m. Hence r is contained in C since α is, and also in C_m since α_m is. But this is impossible, because C and C_m are mutually exclusive.

THEOREM 4. *If S is not a torus and if there is on S a sequence of mutually exclusive circuits of the type $\alpha, \alpha_k, \alpha_{2k}, \cdots$ some pair of which bounds a ring on S, then there exists a generating ring on S, one of whose boundary circuits is α.*

The theorem holds also if S is a torus, provided α is the boundary of a 2-cell on S.

Proof. Suppose that α_{pk} and α_{qk} $(p < q)$ bound a ring. Then α and α_m, $m = (q-p)k$, bound a ring r.

There are several cases to be considered. In the first place, if the rings $r, r_m, r_{2m}, \cdots$ are mutually exclusive, r is the desired generating ring.

Next, suppose that r and r_m are mutually exclusive while r_{am} and r_{bm} $(a < b)$ overlap. Let $h > 1$ be the smallest integer such that r and r_{hm} overlap. The mutually exclusive rings $r, r_m, \cdots r_{(h-1)m}$ together with their boundaries may be considered as constituting a tube the two ends of which overlap on the addition of r_{hm}. Thus a torus is formed which must be identical with S and on which α is non-cell-bounding. This situation is contrary to the hypothesis and the case under consideration is therefore not possible.

Suppose finally that r and r_m overlap. Since α, α_m, and α_{2m} are mutually exclusive, this can happen in only two ways:

(a) One of the two rings is contained in the other.

(b) The two rings have in common a ring s, a proper part of both and bounded by α_m and α_{2m} (i. e., the outer edges of the rings overlap).

Case (b) however is not possible, for the rings s, $r - s$, and $r_m - s$ taken together constitute a torus which must be identical with S and on which α is non-cell-bounding, contrary to hypothesis.

This leaves only case (a). Assume, say, that r contains r_m. Then also

$$(1) \qquad r_m \supset r_{2m} \supset \cdots.$$

Let ρ be the ring $r - (r_m + \alpha_{2m})$. From the relation $\rho \subsetneq r$, we have

$$(2) \qquad \rho \subset r, \qquad \rho_m \subset r_m, \cdots.$$

From (1) and (2), each ring of the sequence $\rho_m, \rho_{2m}, \cdots$ is contained in r_m and therefore fails to intersect ρ. Hence the rings $\rho, \rho_{2m}, \rho_{4m}, \cdots$ are mutually exclusive, for if ρ_{2am} and ρ_{2bm} intersect, so do ρ and $\rho_{(2b-2a)m}$. Therefore, since the boundary of ρ is $\alpha + \alpha_{2m}$, ρ is the desired generating ring. This completes the proof of the theorem.

5. *The connectivity of the regular components.*

THEOREM 5. *A regular component is a region of planar type.*

Let C be a regular component and α an arbitrary circuit in C. We must show (§ 1) that C is separated by α.

Let us assume the contrary and suppose first that α is two-sided. If τ is a small arc crossing α at Q, the two end-points of τ can be joined by a second arc contained in C and intersecting neither α nor τ. Thus we obtain a circuit β contained in C, crossing α at Q, and having no other point in common with α. Clearly, both α and β are non-cell-bounding on S.

Since the points of α are ω-regular, the successive images $\alpha_1, \alpha_2, \cdots$ converge uniformly (§ 3) toward M. Since the closed set $\alpha + \beta$ is at a non-zero distance from M, a positive integer K can be chosen such that α_n fails to meet $\alpha + \beta$ for each $n \geqq K$. The circuits $\alpha_K, \alpha_{2K}, \cdots$ are mutually exclusive. For if α_{pK} and α_{qK} $(p < q)$ intersect, so do α and $\alpha(q - p)_K$ which is impossible by the choice of K. Moreover, these circuits are all non-cell bounding and two-sided, since α is. Hence by the lemma (§ 2) two circuits, say α_{mK} and α_{nK} $(m < n)$ can be chosen which bound a ring on S. Hence α and $\alpha_{(n-m)K}$ bound a ring, say R. Now if we trace a complete circuit on β in the proper sense, we shall enter R by crossing α at Q; but on leaving R, we cannot again cross α and therefore β crosses $\alpha_{(m-n)K}$ which is impossible. This disposes of the case where α is assumed to be two-sided.

If α is one-sided, so are the circuits $\alpha_1, \alpha_2, \cdots$. However, the reasoning above which led to the integer K still applies and we have therefore an infinite sequence $\alpha, \alpha_K, \alpha_{2K}, \cdots$ of mutually exclusive one-sided circuits. But this is impossible, since there can exist on S at most N mutually exclusive one-sided circuits, where N is finite and depends only on the invariants of S. This completes the proof.

We shall say that a regular component of C is *simple*, if every circuit in C is the boundary of a 2-cell on S.

THEOREM 6. *A simple regular component of wandering type is simply connected.*

Let C be such a component, and Γ its boundary. Also let α be an

arbitrary circuit in C and A a 2-cell with boundary α. To demonstrate the theorem it will be sufficient to show that A is contained in C (§ 1).

Suppose this were not the case. Then A contains points of Γ and each of these is either irregular or non-wandering. It follows then from theorem 1 that A cannot be wandering. Let h be a positive integer such that A and A_h intersect. Since α and α_h fail to intersect, one of the following situations holds:

(a) One of the cells A, A_h together with its boundary is contained in the other.

(b) The part common to A and A_h is a ring bounded by $\alpha + \alpha_h$.

We shall first consider the case (a). Suppose for example that $A \supset A_h + \alpha_h$. Then we have $A \supset A_h \supset A_{2h} \supset \cdots$, and if we call r the ring $A - (A_h + \alpha_h)$, the images $r, r_h, r_{2h}, \cdots$ are mutually exclusive. The boundary of r is $\alpha + \alpha_h$ and hence r is a generating ring, one of whose boundary circuits, namely α, is in C. This is impossible by theorem 3.

(b) In this case, A and A_h together cover S, and S is equivalent to a sphere. The cells A_h and A_{2h} overlap just as A and A_h do, and it is readily seen that the cells A and A_{2h} therefore overlap in such a way that one is contained in the other. Thus we are reduced to the case (a) with h replaced by $2h$. This completes the proof.

THEOREM 7. *If S is not a torus, every regular component of wandering type on S is simple.*

For, let C be such a component. If C is not simple, a circuit α can be chosen in C which fails to bound a cell on S. The mutually exclusive non-cell-bounding circuits $\alpha, \alpha_1, \alpha_2, \cdots$ are two-sided since C is of planar type (theorem 5). Hence by the lemma (§ 2), there exist integers a and b $(0 < a < b)$ such that α_a and α_b bound a ring on S. Hence by theorem 4 there exists a generating ring one of whose boundary circuits is α. But since α is in C, we have a contradiction to theorem 3.

THEOREM 8. *On a surface which is not a torus, every regular component of wandering type is simply connected.*

This is a consequence of theorems 6 and 7.

THEOREM 9. *A regular component which contains at least one boundary circuit of a generating ring r is doubly connected.*

We shall show in fact that C is identical with the doubly connected limit region

$$R = r' + r'_h + \cdots + r'_{-h} + \cdots$$

where h is the order of r (§ 4).

Let the boundary of r be $\alpha + \alpha_h$ where one of these circuits, which we may assume to be α, is contained in C. Also let the boundary of R be $\rho^1 + \rho^2$ where ρ^1 and ρ^2 are respectively the continua toward which the sequences r_h, $r_{2h}, \cdots$ and r_{-h}, $r_{-2h}, \cdots$ converge.

Assuming definitely that C is ω-regular, we have immediately that $\rho^1 \subseteq M$. For, an arbitrary point Q is a limit point of a sequence Q^h, $Q^{2h}, \cdots$ of points suitably chosen on the respective circuits α_h, $\alpha_{2h}, \cdots$. These circuits however converge uniformly toward M (§ 3) and it follows that Q is a point of M.

Next, the points of ρ^2 which are not contained in M, if any exist, are ω-irregular. For, let P be such a point and σ a neighborhood of P chosen so small that it contains no points of M. Now σR converges toward ρ^1 in iteration of T_h, whereas the points P, P_h, $P_{2h}, \cdots$ are all on ρ^2. Hence the successive images σ, $\sigma_h, \cdots$ tend to stretch across the region R, and must therefore, from a certain rank on, all intersect the circuit α which is at a non-zero distance from M. Hence σ does not converge uniformly toward M on iteration of T and P is therefore ω-irregular.

It follows that although C contains points of R, for example α, C contains no boundary points of R. Hence $C \subseteq R$.

We have finally to show that $R \subseteq C$. It follows from theorem 1 that the points of r' and hence of every image of r' are ω-regular. Thus R consists only of ω-regular points and since some points of R are in C (e. g. α), R is entirely contained in C. This completes the proof.

THEOREM 10. *Every regular component is at most doubly connected.*

Let C be a regular component and α an arbitrary circuit in C; by theorem 5, α is two-sided. If α is the boundary of a 2-cell contained in C, then C is simply connected (§ 1); in the contrary case C is doubly connected as we proceed to prove.

Suppose that α is not the boundary of a cell contained in C. Since α consists of ω regular points and is at a non-zero distance from M, a positive [illegible]

contained in C, it contains points which are not ω-regular, namely, boundary points of C. Hence by theorem 1, the cells A, A_N, $A_{2N}, \cdots$ cannot be mutually exclusive. Suppose then that A_{aN} and A_{bN} intersect. Then one of these cells together with its boundary is contained in the other, or else the part common to A_{aN} and A_{bN} is a ring. In either case $\alpha_{aN} + \alpha_{bN}$ is the boundary of a ring on S. Since α is cell-bounding, theorem 4 is applicable even when S is a torus with the result that there exists a generating ring on S one of whose boundary circuits is α. Hence by theorem 9, C is doubly connected.

There remains the case in which α is non-cell-bounding. Then α_N, $\alpha_{2N}, \cdots$ are non-cell-bounding and also two-sided. Hence by the lemma (§ 2) a pair of these circuits can be chosen which bounds a ring on S. If S is not a torus, theorem 4 is again applicable and then, as above, theorem 9, with the result that C is doubly connected.

In the excepted case, α may be assumed to bound a cell on S, as we shall show, and then we are reduced to the case already treated. For assume S to be a torus. Then *if every circuit α which fails to bound a cell in C, also fails to bound a cell on S, C is at most doubly connected.* To prove this, suppose that C is neither simply nor doubly connected. Then there exists in C (§ 1) a pair of non-intersecting circuits β and γ neither of which is the boundary of a cell in C, and such that $\beta + \gamma$ is not the boundary of a ring in C. By assumption, β and γ are non-cell-bounding on S and hence (§ 2) separate S into two rings, r^1 and r^2, and both of these contain boundary points of C. Let β and γ be joined by a simple arc in C. A suitably chosen sub-arc τ will connect β and γ and will be entirely contained in r^1 or r^2, say r^1. On making a cut along τ, r^1 is reduced to a simply connected region r^* which contains boundary points of C and whose boundary is contained in C. On tracing a curve in r^* close to its boundary we obtain a circuit which fails to bound a cell in C, but bounds a cell on S. This contradiction completes the proof of the theorem.

6. *Transformations of a Torus.* We have seen that regular components of wandering type are simply connected if S is not a torus. On a torus, however, such components may actually be doubly connected. For example let S be a circular torus and θ and ϕ its angular co-ordinates. It is possible to define a transformation t of the circle $\phi = 0$ into itself which admits a regular component (arc segment) of wandering type.† If each circle ϕ

† For a detailed description of such a transformation, see Birkoff, "Quelques Théorèmes sur le Mouvement des Systèmes dynamiques," *Bulletin de la Société Mathématiques de France*, Vol. 40, p. 16.

$=$ const. be subjected to a transformation congruent to t, we have a transformation T of the desired type. The regular components are wandering rings bounded by circles of the type $\theta =$ const.

Not only is the existence of regular components of the type under consideration restricted to surfaces of special type, but their existence implies a certain speciality in the structure of T itself as we shall show. In the remaining paragraphs we shall assume that S is a torus and that T admits regular components of doubly connected wandering type. We shall denote by D an arbitrarily chosen one of these components.

I. *The closed set $E = S - D$ contains no invariant or periodic points of continua.*

Suppose on the contrary that E does contain a periodic continuum, ξ say, of order h. By theorem 6, D is non-simple and hence there exists in D a circuit α which is non-cell-bounding; α fails to meet ξ, and ξ being invariant under T_h, each image $\alpha_h, \alpha_{2h}, \cdots$ fails to meet ξ. The mutually exclusive circuits α and α_h separate S into two rings (§ 2) ρ and σ, one of which, say ρ, fails to contain points of ξ; hence each of the rings $\rho_h, \rho_{2h}, \cdots$ fails to contain points of ξ.

The rings ρ and ρ_h are mutually exclusive. For suppose they intersect; then one is contained in the other, or else the edges along α and α_{2h} respectively overlap. But this latter situation is impossible, since ρ and ρ_h taken together would cover S, whereas neither ρ nor ρ_h contains points of ξ. In the former case, we are led by the reasoning in the proof of theorem 4, case (a), to the existence of a generating ring on S, one of whose boundary circuits is α. This contradicts theorem 3.

Thus ρ and ρ_h fail to intersect. Moreover, ρ is intersected by none of its images $\rho_h, \rho_{2h}, \cdots$. For on adding successive rings to ρ, we build a tube and at no stage will the two ends overlap for otherwise a finite number of rings would cover S, whereas none of them contains points of ξ. Hence ρ is a generating ring which again contradicts theorem 3. This completes the proof.

Since D itself contains no invariant or periodic points, we have as a corollary that *there are no invariant or periodic points whatever on S.*

II. *T admits no regular components of wandering type.*

For suppose the regular component E is periodic. Since E is at most doubly connected, its boundary consists of two continua (which may intersect). These continua are contained in $S - D$ and both are invariant or periodic, which is impossible by I.

Let α continue to represent an arbitrary non-cell-bounding circuit in C and let Q^n be an arbitrarily chosen point on α_n $(n = 1, 2, \cdots)$. A limit point of the sequence $Q^1, Q^2, \cdots$ belongs necessarily to M on account of the uniform manner in which α converges toward M on indefinite iteration of T (§ 3). Let L be the totality of the limit points obtained from all possible sequences of the type $Q^1, Q^2, \cdots$; L is clearly a closed invariant subset of M. Moreover, L is independent of the choice of D and α; in fact *L is identical with M* as will presently appear.

Let E be an arbitrary component of the open set $S - L$; clearly E, like D, is either wandering or periodic.

A 2-cell B on S whose boundary β is in E cannot contain boundary points of E. For suppose B contains a boundary point Q. Let $Q^1, Q^2, \cdots$ be a point sequence of points which has Q for a limit point and which is of the type which determines L. Infinitely many points of this sequence are in B, and the corresponding circuits of the sequence $\alpha_1, \alpha_2, \cdots$, since they are non-cell-bounding, all intersect β. Hence on β there is at least one point of L, which is impossible.

By a similar argument, *L is perfect.* For suppose that Q is a point of L and B an arbitrary 2-cell containing Q. Then on the boundary of B there is at least one point of L, and hence Q is a limit point of L.

The set L is nowhere dense. For suppose that L contains an inner point P. Let T be that maximal connected subset of inner points of L which contains P. Since F contains only non-wandering points, it is intersected by some of its images, one of which, say, is F_h. But since F is maximal, we have $F \equiv F_h$. Thus F together with its boundary is a periodic continuum in $S - D$ which is impossible by I.

Since the boundary of E is in M, it follows by theorem 2, § 3, *that E is an ω- and α-regular component.*

We are now prepared to show that L is identical with M. Since $L \subseteq M$, it remains only to show that $M \subseteq L$. Suppose on the contrary that there is a point P of M which is not in L. Then P is contained in one of the components of $S - L$, say in H. But this is impossible, for by the preceding paragraph, H contains no non-wandering points.

Finally, we shall show that *each regular component is identical with a component of $S - L$.*

Let C be an arbitrary regular component. Then there exists a component F of $S - M$ such that $C \subseteq F$. But since $M \equiv L$, F is also a component of $S - L$ and therefore is a regular component, as we have shown. It follows that C cannot be a proper part of F and hence $C \equiv F$.

The preceding result is interesting for the following reason: since C is identical with a component of $S - L$, the boundary of C consists only of non-wandering points. Hence *T admits no irregular points.*

We now summarize the results of this section in the following

THEOREM. *If T admits a regular component of doubly connected wandering type, then* (a) *S is a torus;* (b) *M is a perfect nowhere dense set which contains no invariant or periodic continua;* (c) *T admits no invariant or periodic points;* (d) *T admits no irregular points;* (e) *the regular components of T are both ω- and α-regular and there are none of periodic type.*

A surface transformation T with these properties is in a sense analogous to a transformation of a simple closed curve with rotation number [*] incommensurable with 2π and non-wandering points constituting a perfect nowhere dense set. It is possible in fact to attach a unique rotation number to T itself, as we shall indicate. Let α be an arbitrary non-cell-bounding circuit in C. Any two circuits of the sequence

$$\cdots, \alpha_{-1}, \alpha, \alpha_1, \cdots$$

separate S into two rings and these circuits can therefore be ordered cyclically on S. Now let $\delta^1, \delta^2, \cdots$ be an infinite sequence of mutually exclusive arcs on a circle σ, so chosen that the set $\sigma - \Sigma\delta^i$ is perfect and nowhere dense. By establishing a $(1, 1)$ correspondence H between the circuits α_i and the arcs δ^j in such a way as to preserve order, a transformation of the set $\Sigma\delta^i$ into itself is induced by T, and this transformation extends by continuity to the whole of σ; its rotation number λ is incommensurable with 2π. Moreover λ is independent of the choice of the correspondence H and is therefore characteristic of T. We shall not, however, carry this discussion into further detail.

BARNARD COLLEGE,
COLUMBIA UNIVERSITY.

Polygenic Functions of the Dual Variable $w = u + jv$.

By Edna E. Kramer.

It is the purpose of this paper to develop a theory *analogous*, and, roughly speaking, *dual* to Kasner's recent theory of polygenic functions of z.* There seems to be no general principle of transference whereby the properties of the functions here considered could be directly inferred from the corresponding facts in the original theory. A summary of some of the results obtained below will serve to illustrate points of similarity and difference in the two theories.

The first derivative of a polygenic function of z is represented by a congruence of *circles*. For functions of the dual variable we have a congruence of directed circles, or *cycles*.† In the matter of the method of generation of these curves, the theories diverge. Kasner has found it appropriate, with this question in view, to speak of a congruence of *clocks*. It will be shown below that our derivative cycles are described in an entirely different way. There is also some variation in the nature of the specialization of the congruences. For the second derivative we have as corresponding pictures, a *line* and a *cycle*, a curve of *eighth* order and a directed curve of *sixth* order.

1. *The First Derivative.* Just as the *conformal* group is represented by the monogenic functions of the ordinary complex variable $z = x + iy$, the *equilong* group ‡ is represented by the monogenic functions of the dual variable

$$w = u + jv, \text{ where } j^2 = 0,$$
$$u = tg(\theta/2), \qquad v = (\rho/2)\sec^2(\theta/2),$$

---

* E. Kasner, "Theory of Polygenic Functions," *Science*, Vol. 66 (1927), pp. 581-582, and *Proceedings of the National Academy of Sciences*, Vol. 14 (1928), pp. 75-82; "Second Derivative of a Polygenic Function," *Transactions of the American Mathematical Society*, October 1928; "Note on the Derivative Circular Congruence of a Polygenic Function," *Bulletin of the American Mathematical Society*, Vol. 34 (1928), pp. 561-565; also L. Hofmann and E. Kasner, "Homographic Circles or Clocks, with an Appendix on Polygenic Functions," by E. Kasner, the same *Bulletin*, Vol. 34, pp. 495-503.

† We shall deal, in general. with *directed* curves, and refer to a directed circle as a *cycle*, and to a directed line as a *ray*.

‡ G. Scheffers, "Isogonalkurven, Aquitangentialkurven, und Komplexe Zahlen," *Mathematische Annalen*, 1905.

and θ and ρ are oriented Hessian line co-ordinates. (If t is the perpendicular from the origin O upon a given ray s, oriented relative to s as the positive X to the positive Y axis, then ρ is equal to the distance on t from O to the intersection of s and t, and is positive or negative according as s indicates positive or negative rotation about O, and θ is the angle made by t with a fixed ray through O.)

$W = \phi(u, v) + j\psi(u, v)$ is a *monogenic* function of w, or represents an equilong transformation if

$$\phi_v = 0, \qquad \phi_u = \psi_v, \tag{1}$$

which are the analogue of the Cauchy-Riemann equations for functions of z.

We give a definition of *polygenic* functions of w analogous to that which Kasner has given for polygenic functions of z.

We shall call those functions $W = \phi(u, v) + j\psi(u, v)$ polygenic, in which ϕ and ψ are continuous, and have continuous partial derivatives in the region considered, but do not necessarily satisfy the equations (1).

The limit of the ratio $\Delta W/\Delta w$ depends, in general, not only on the ray $u + jv$, but also on the point in the ray through which the neighboring ray passes as it approaches coincidence with the given ray. Thus the derivative dW/dw has, in general, infinitely many values for a given ray, unless equations (1) are satisfied, when there is a unique value.

Now

$$\begin{aligned} dW/dw &= [\phi_u + j\psi_u + v'(\phi_v + j\psi_v)]/(1 + jv') \\ &= \phi_u + v'\phi_v + j[\psi_u + v'(\psi_v - \phi_u) - v'^2\phi_v]. \end{aligned}$$

Let $\alpha + j\beta = dW/dw$. Then

$$\alpha = \phi_u + v'\phi_v, \qquad \beta = \psi_u + v'(\psi_v - \phi_u) - v'^2\phi_v'. \tag{2}$$

Eliminating v' between these two equations, we have

$$\beta = (\phi_v\psi_u - \phi_u\psi_v)/\phi_v + [(\psi_v + \phi_u)/\phi_v]\alpha - (1/\phi_v)\alpha^2, \tag{3}$$

the equation of a *cycle*.* If (A, B) is the center, and C the radius

$$\begin{aligned} A &= (\phi_v\psi_u - \phi_u\psi_v + 1)/\phi_v, \\ B &= (\phi_u + \psi_v)/\phi_v, \\ C &= (\phi_v\psi_u - \phi_u\psi_v - 1)/\phi_v. \end{aligned} \tag{4}$$

To each ray of the plane corresponds, in general, a cycle. To the ∞^2 rays correspond, in general, ∞^2 cycles.

The above work breaks down only if $\phi_v = 0$.

* The equation of a cycle in the coordinate system used here is $v = a + bu + cu^2$.

If $\phi_v = 0$ and $\phi_u = \psi_v$, that is, if the function is monogenic, we have the ray

$$\alpha = \phi_u, \qquad \beta = \psi_u.$$

Otherwise we have the two points at infinity whose equation is

$$\phi_u \psi_v - (\psi_v + \phi_u)\alpha + \alpha^2 = 0.$$

If $\phi_u = -\psi_v$, that is, if the transformation represented by W is indirect equilong, these are the points at infinity $\alpha = \pm \phi_u$.

These conditions may be satisfied for special rays only. For such rays, then, the derivative cycle will degenerate into a ray or a point at infinity. If the conditions are identically satisfied, this is true for all rays.

The following theorems deal with special cases:

I. The congruence of cycles will become a congruence of points if

$$\phi_u \psi_v - \phi_v \psi_u = -1;$$

that is, if the Jacobian of ϕ and ψ is equal to -1.

II. The centers of the cycles will lie on the y-axis if the Jacobian of ϕ and ψ is equal to 1.

III. The centers of the cycles will lie on the x-axis if $\phi_u = -\psi_v$, so that $\phi = \chi_v$, $\psi = -\chi_u$; that is, $W = \chi_v - j\chi_u$, where χ is an arbitrary function of u and v.

IV. The ∞^2 cycles will reduce to ∞^1 with a common center at the origin if both the conditions given in II and III are satisfied.

V. If we require that the congruence be such as to include the reverse cycle to each one of the set (and hence to become a congruence of circles) we find that this cannot occur unless the congruence is one of points, which is the condition in I.

VI. Two polygenic functions $\Phi + j\Psi$ and $\phi + j\psi$ are related as follows, if for each ray of the plane the derivative cycles for the two functions are opposite cycles of the same circle:

$$(\Phi_u + \Psi_v)/\Phi_v = (\phi_u + \psi_v)/\phi_v,$$
$$(\Phi_v \Psi_u - \Phi_u \Psi_v)/\Phi_v = 1/\phi_v,$$
$$1/\Phi_v = (\phi_v \psi_u - \phi_u \psi_v)/\phi_v.$$

The last two equations are equivalent to the condition that the Jacobians be reciprocal.

VII. If the cycles are to be tangent to $w = 0$, the Jacobian of ϕ and ψ must vanish.

VIII. If the cycles are to be tangent to any fixed ray (a, b) the function must satisfy

$$\phi_v \psi_u - \phi_u \psi_v + a(\psi_v + \phi_u) - a^2 = b\phi_v.$$

IX. If the ∞^2 cycles are to reduce to the ∞^1 cycles tangent to $w = 0$ at the origin, the Jacobian must vanish, and in addition

$$\phi_u = -\psi_v.$$

X. The condition that all the centers lie at a fixed point is

$$\phi_v \psi_u - \phi_u \psi_v + 1 = a\phi_v, \qquad \psi_v + \phi_u = b\phi_v.$$

XI. The congruence will reduce to a single cycle if

$$\phi_v \psi_u - \phi_u \psi_v - 1 = a\phi_v, \quad \phi_v \psi_u - \phi_u \psi_v + 1 = b\phi_v, \quad \phi_u + \psi_v = c\phi_v,$$

By Fiedler's method of projection * a congruence of cycles corresponds to a real surface. Then the derivative of a polygenic function can be represented by a surface, which will be a plane when

$$(a + b)(\phi_v \psi_u - \phi_u \psi_v) + a - b + c(\phi_u + \psi_v) + d\phi_v = 0,$$

where a, b, c, d are constants.

We reword the above theorems as follows:

I′. The derivative surface of a polygenic function will be the plane $z = 0$ if the Jacobian of ϕ and ψ is equal to -1.

II′. The derivative surface will be the plane $x = 0$ if the Jacobian is equal to 1.

III′. The surface will be the plane $y = 0$ if $\phi_u = -\psi_v$.

IV′. The derivative will be represented by the Y axis if both conditions given in II′ and III′ are satisfied.

VI′. If the derivative surfaces of two polygenic functions are to be obtainable from one another by reflection in the XY plane, the conditions of VI must be satisfied.

VII′. The plane $x + z = 0$ represents the derivative of a polygenic function with vanishing Jacobian.

---

* Fiedler, *Zyklographie*. If A, B are the Cartesian coordinates of the center and C the radius of a cycle, then $x = A$, $y = B$, $z = C$.

IX′. If in addition to the vanishing of the Jacobian, we have $\phi_u = -\psi_v$ the derivative is represented by the line

$$y = 0, \qquad z = -x.$$

X′. If the conditions of X are satisfied, the derivative is represented by the line

$$x = a, \qquad y = b.$$

XI′. The conditions of XI will give a point as the picture of the derivative.

We shall next examine the method of description of the derivative cycle.

From equation (2) for α it is evident that the change in α is proportional to the change in v', the ratio of the rates of change being equal to ϕ_v.

Let (θ, ρ) be Hessian co-ordinates of the ray (α, β). As v' varies, that is, as the point of approach moves along the given ray in the (u, v) plane, the point of contact of the ray (α, β) moves along the derivative cycle, and the ratio of the rates of the two points is expressed by

$$(1/C)\, dv'/d\theta,$$

where C has the value given above.

If we assume that the distribution of points on the ray in the (u, v) plane is uniform, then the absolute value of this ratio may be called the *density* of points on the derivative cycle. Representing this quantity by δ we have

$$\delta = \sec^2 \tfrac{1}{2}\theta/2 \mid \phi_v\psi_u - \phi_u\psi_v - 1 \mid . \tag{5}$$

Thus the distribution of points on the derivative cycle will be uniform only if θ is constant, which will occur only when the derivative cycle degenerates into a single ray or two points at infinity, (that is, if $\phi_v = 0$). If

$$\phi_v \neq 0 \quad \text{and} \quad \phi_u\psi_v - \phi_v\psi_u \neq -1,$$

then as v' increases from $-\infty$ to $-\phi_u/\phi_v$ to $+\infty$, the point of contact on the derivative cycle moves from the position $\theta = -\pi$ to $\theta = 0$ to $\theta = \pi$, and δ varies from ∞ to $(1/2) \mid \phi_v\psi_u - \phi_u\psi_v - 1 \mid$ to ∞, assuming equal values at points of the derivative cycle symmetrical with respect to the horizontal diameter.

2. *The Second Derivative.* Let $\sigma = \sigma_1 + j\sigma_2$ represent the second derivative of the polygenic function $W = \phi + j\psi$. Then

$$\sigma = (W_{uu} + 2v'W_{uv} + v'^2 W_{vv})/(1 + jv')^2 + v''(W_v - jW_u)/(1 + jv')^3, \tag{6}$$

and hence

$$
(7) \qquad
\begin{aligned}
\sigma_1 &= \phi_{uu} + 2v'\phi_{uv} + v'^2\phi_{vv} + v''\phi_v, \\
\sigma_2 &= \psi_{uu} + 2v'(\psi_{uv} - \phi_{uu}) + v'^2(\psi_{vv} - 4\phi_{uv}) \\
&\qquad - 2v'^3\phi_{vv} + v''(\psi_v - \phi_u - 3v'\phi_v).
\end{aligned}
$$

The second derivative of a polygenic function will thus depend, in general, not only on the ray for which the derivative is taken, but on v' and v'' as well.

It will be independent of v'' only when

$$\phi_v = 0 \quad \text{and} \quad \phi_u = \psi_v,$$

that is, only when the function is monogenic.

The second derivative can be expressed as a function of v' and r, the radius of curvature, by the substitution

$$v'' = 2(r + uv' - v)/(1 + u^2)$$

in the above expression. We have

$$
(8) \qquad
\begin{aligned}
\sigma_1 &= \phi_{uu} - [2\phi_{vv}/(1 + u^2)] \\
&\quad + 2\{\phi_{uv} + [u\phi_v/(1 + u^2)]\}v' + \phi_{vv}v'^2 + [2\phi_v r/(1 + u^2)], \\
\sigma_2 &= \psi_{uu} + [2(\phi_u - \psi_v)v/(1 + u^2)] \\
&\quad + 2\{\psi_{uv} - \phi_{uu} + [u(\psi_v - \phi_u) + 3\phi_v v/(1 + u^2)]\}v' \\
&\quad + (\psi_{vv} - 4\phi_{uv} - [6u\phi_v/(1 + u^2)]v'^2 \\
&\quad - 2\phi_{vv}v'^3 + [2(\psi_v - \phi_u - 3v'\phi_v)r/(1 + u^2)].
\end{aligned}
$$

If we fix v' and allow r to vary, that is, if we fix a ray in the (u, v) plane, and a point on the ray, and allow a neighboring ray to approach the given ray along different tangent cycles, we find that the values which the second derivative assumes may be plotted as a *cycle*. For, if we eliminate r between the equations (8) we obtain

$$(9) \qquad \sigma_2 = a + b\sigma_1,$$

where a and b are functions of u, v and the partial derivatives.

This is a cycle tangent to the reverse Y axis.

From the equations (8) we see that the change in σ_1 is proportional to the change in r. As r varies the cycle (9) is generated in a manner similar to that explained for the first derivative cycle above.

Now if we let v' vary we have ∞^1 cycles tangent to the reverse Y axis whose centers lie on a curve

If in (8) we hold r constant and allow v' to vary, that is, if we consider the ∞^1 values of the second derivative for a given ray as this ray is approached along elements of the same curvature, then we have as the picture of the second derivative the curve represented by (8) if we regard v' as parameter.

In general, this curve (non-oriented) is one of the *sixth* order. If $\phi_{vv} = 0$, this curve becomes a *cycle.*

3. *Linear Fractional Polygenic Functions.* By a linear fractional polygenic transformation we shall mean one of the set of ∞^{10} transformations

$$W = (Aw + B\bar{w} + C)/Dw + E\bar{w} + F), \tag{10}$$

where the constants are dual numbers.

This set of transformations is of interest, since it includes the Laguerre Group as a sub-group. It has also the sub-group

(11)

$$\begin{aligned} U &= (au + bv + c)/(du + ev + f), \\ V &= (gu + hv + k)/(du + ev + f). \end{aligned}$$

It will reduce to the former group if

$$B = E = 0 \quad \text{or} \quad A = D = 0.$$

It will reduce to the latter if

$$\bar{D}/E = \bar{E}/D = \bar{F}/F.$$

The group (11) will transform cycles tangent to the reverse y-axis into cycles, but will transform other cycles into curves of higher order.

The set (10) also includes the group

$$W = aw + b\bar{w} + c, \tag{12}$$

which converts cycles into cycles.

All invariants under (12) are combinations of

$$v^{(n)}(v'')^{n-3}/(v''')^{n-2}, \qquad \text{where } n \geqq 4.$$

Further discussion of these transformations and invariants will be published elsewhere.

Trinomial Curves and Monomial Groups.*

By R. M. Winger.

---

1. *Introduction.* Maschke † has considered a particular class of monomial collineation groups in the ternary domain, namely groups generated by the two substitutions

$$x' = y, \qquad y' = z, \qquad z' = x,$$
$$x' = ax, \qquad y' = by, \qquad z' = cz,$$

where a, b, c are roots of unity and $abc = 1$. These restrictions exclude even monomial groups requiring more than two generators, as well as those with two generators one of which is a homology whose period is a multiple of 3. Furthermore any of Maschke's groups can be enlarged by adding as a generator a homology (of period greater than 2) having a vertex and opposite side of the fixed triangle for center and axis. Skinner ‡ has extended the study to a larger class of monomial groups, viz. those containing only elements of determinant ± 1.

Both authors, however, exclude the interesting groups that leave invariant the trinomial curves $x^n + y^n + z^n = 0$. These curves have received attention at the hands of numerous writers.§ But the group aspect of the curves has been neglected except for a few particular cases. From this point of view Berzolari ¶ has treated the projective lemniscate ($n = -2$) whose associated group is the ternary octahedral G_{24}. The quartic case has been studied by Dyck.‖ Snyder ** notices the case $n = 5$ in his study of quintic curves invariant under linear transformations. Tappan †† gives the generators of the group for the sextic, while Musselman ‡‡ discusses the group and curve briefly. Finally the author ¶¶ has discussed in considerable detail

---

* Read before the American Mathematical Society, December 28, 1927.

† *American Journal of Mathematics*, Vol. 17 (1895), p. 168.

‡ *Ibid.*, Vol. 25 (1903), p. 17.

§ For references, see Loria, *Spezielle Algebraische und Transzendente Ebenen Kurven*, Vol. 1, p. 328 ff.

¶ *Istituto Lombardo, Rendiconti*, Ser. 2, Vol. 37 (1904), pp. 277, 301.

‖ *Mathematische Annalen*, Vol. 17 (1880), p. 510.

** *American Journal of Mathematics*, Vol. 30 (1908), p. 7.

†† *Ibid.*, Vol. 37 (1915), p. 320.

‡‡ *Ibid.*, Vol. 49 (1927), p. 355.

¶¶ *Tôhoku Mathematical Journal*, Vol. 29 (1928), pp. 376-400. Muth, *Dissertation*, Giessen (1890), devotes some space to the equianharmonic cubic, $n = 3$.

the cases $n = \pm 3$ which are invariants of the monomial subgroup G_{54} of the ternary Hesse group. The generalization of the principal results of that paper is the present objective.

The purpose of the paper is thus twofold: the analysis of the group and an investigation of trinomial curves for an arbitrary integral value of n. The invariant configuration of the group is described and the relation of invariant curves thereto is pictured. The complete system of invariants of the group is obtained and several constructions of conjugate sets of points on individual curves of the systems are explained. A quadratic inversion is introduced which proves to be an effective auxilliary weapon to the collineation group. In conclusion some generalizations to n-space are sketched.

Broadly speaking we find that the geometry of the group and the invariant curves differs according as n is odd or even and again according to the classification of n in the residue system with respect to the modulus 3. This provides a natural basis for subdividing the curves into species.

2. *The Structure of the Group.* If we consider the substitutions on the variables which leave unaltered the curve

$$x^n + y^n + z^n = 0, \qquad n \text{ a positive integer},$$

they must be either permutations of the letters or permutations of the letters multiplied by an nth root of unity.* All such substitutions give rise to collineations which can be generated by the following set

$$\begin{array}{rcccc} & S & T & R & W \\ x' = & \epsilon x & z & y & \epsilon x \\ y' = & \epsilon^{-1} y & x & x & y \\ z' = & z & y & z & z, \end{array} \qquad \epsilon = e^{2\pi i/n}.$$

For T and R in combination effect all permutations of the letters, while products of powers of S and W have the effect of multiplying the letters by the nth roots of unity in all possible ways.

We suppose throughout that the variables are homogeneous projective co-ordinates, that the transformed variables are primed, and that multiplication if from left to right. The vertices of the triangle of reference opposite the sides x, y, z, respectively, we shall denote by u, v, w. We shall at pleasure refer to the individual collineations of the group as "elements." An element of period two, i. e. a harmonic homology, we shall call a "reflexion."

We observe first that *the order of the group is* $6n^2$. For there are n^2

* There is an exception in the case of the conic which may admit still other types of substitutions.

projectively distinct collineations which do not disturb the order of the variables, viz.

$$(1) \qquad x' = x, \quad y' = \epsilon^i y, \quad z' = \epsilon^j z, \qquad i, j = 1, 2, \cdots, n.$$

These n^2 elements form a subgroup whose generators are S and W. Now T and R generate a dihedral subgroup of order 6, the products of whose elements by the n^2 elements (1) aggregate a total of the $6n^2$ projectively distinct collineations. It follows that the elements of the group fall into six sets of n^2 each, corresponding to the six permutations of the variables. Or,

The group is in (n^2, 1) *isomorphism with* (a) *the permutation group on three letters,* (b) *the dihedral group of order* 6 *generated by* T *and* R.

The dihedral G_6 is thus a factor group of the whole group. The invariant subgroup corresponding to the identical element of the factor group is that generated by S and W. This group is also Abelian since it comprises all those elements which do not permute the letters. Henceforth we shall denote the Abelian invariant subgroup, consisting of the elements (1), by G. Further by GT, for example, we shall mean the n^2 elements formed by multiplying each element of G by T. Then the six sets of n^2 elements of our group may be represented symbolically as follows:

$$(2) \qquad \begin{array}{ccc} G & GT & GT^2 \\ GR & GTR & GT^2R \end{array}$$

or

$$(3) \qquad \begin{array}{ccc} S^iW^k & S^iW^kT & S^iW^kT^2 \\ S^iW^kR & S^iW^kTR & S^iW^kT^2R, \end{array} \qquad (i, k = 1, \cdots, n).$$

To find the element of the factor group to which the n^2 elements represented by any symbol in (2) correspond, we have merely to replace the "G" in that symbol by 1. Or if we replace by 1 the multipliers of any collineation of the entire group we shall obtain that element of the factor group to which the collineation corresponds.

The triangle of reference obviously is fixed under the group: all the elements of the subgroup G leave each vertex fixed: the elements GT and GT^2 permute the vertices cyclically while the remaining elements leave one

These homologies,—and indeed all elements of G, excepting the identity, are of period n if n is prime. When n is even there is one reflexion in each set of $n-1$.

Again, S and R generate a dihedral subgroup of order $2n$. The n reflexions of this group belong to the set GR. Their centers lie on the z-axis and their axes are concurrent at the opposite vertex. Therefore,

The group includes three dihedral subgroups of order $2n$, each having a vertex and opposite side of the fixed triangle as invariant point-line. Further the group contains $3n$ reflexions whose centers lie in sets of n on the sides of the fixed triangle and whose axes, in sets of n, meet at the vertices.

There is a difference in the dihedral groups according as n is odd or even. When n is odd, the axes of reflexion are conjugate and the centers do not lie on the axes. On the other hand when n is even, the axes divide into two conjugate sets, say ξ_1, and ξ_2 of $n/2$. And each center lies on an axis: the centers associated with ξ_i lie on the axes ξ_j when $n/2$ is odd, and on ξ_i when $n/2$ is even. In any case the equation of the axes of reflexion is

$$(y^n-z^n)(z^n-x^n)(x^n-y^n)=0$$

and the co-ordinates of the centers are

$$(0, \epsilon^i, -1), \quad (\epsilon^i, -1, 0), \quad (-1, 0, \epsilon^i).$$

The multiple isomorphism implies the existence of several subgroups, corresponding to the subgroups of the factor group. Thus the elements of the first row of the table (2) (or (3)) constitute a second invariant subgroup, generated by S, T, W, corresponding to the cyclic G_3 of the factor group. This group is of order $3n^2$ and comprises those elements which effect the cyclic advance (xyz) of the variables, multiplied by roots of unity. It is easily verified that the elements of this group, not contained in G, are all of period three and that the square of each element GT is an element GT^2 and vice versa. Hence

The group contains $2n^2$ elements of period 3, which belong in n^2 cyclic subgroups G_3.

These subgroups, as we shall prove in the next section, are all conjugate when n is not a multiple of 3. When n is a multiple of 3, they form two conjugate sets, containing respectively one-third and two-thirds of the groups.

We have now accounted for half the elements of the group. We have also noted $3n$ reflexions belonging to the other half GT^jR. The remaining elements of this second-half are all of even period since they correspond to

elements of period two in the factor group, but the precise periods depend on the factorability of n.

The elements corresponding to each of the reflexions of the factor group, together with the elements of G, form a subgroup of order $2n^2$. Thus the elements GR, corresponding to R, are $S^i W^k R$, $i, k = 1, \cdots, n$. The square of each of these is some homology of G. For example the collineation $S^i W^k R$ and its square are respectively

$$(4) \qquad x' = \epsilon^{-i} y, \qquad y' = \epsilon^{i+k} x, \qquad z' = z.$$

$$(5) \qquad x' = \epsilon^{k} x, \qquad y' = \epsilon^{k} y, \qquad z' = z.$$

When $k = n$, (4) represents the $3n$ reflexions of the dihedral G_{2n} generated by S and R.

When $k = 1$, the homology (5) is of period n whence (4) is of period $2n$. If now p of the nth roots of unity are primitive, then the number of collineations of period $2n$ in GR is np, for with each value of i $(= 1, \cdots, n)$ can be combined p values of k. Each collineation of period $2n$ generates a cyclic subgroup G_{2n}, of order $2n$. The n even powers of a generator of such a subgroup are homologies belonging to G, but the odd powers are all included in the set GR. The number of cyclic G_{2n}'s, however, depends on the nature of n.

n, odd. Each G_{2n} contains p elements of period $2n$, since there are p primitive $2n$-th roots of unity. Accordingly there are n cyclic G_{2n}'s generated by the elements of period $2n$ in GR. The odd powers of these generators are all distinct and comprise all the elements GR. Similar statements hold for the sets GTR and GT^2R. Hence

The group contains $3n$ cyclic subgroups G_{2n}, of order $2n$ when n is odd. Half the elements of these groups are homologies (including the identity) from G while the other half comprise all the $3n^2$ elements corresponding to the elements of period two in the factor group.

The nth power of any generator of each of these G_{2n}'s is one of the $3n$ reflexions.

n, even. Since there are now $2p$ primitive $2n$-th roots, each cyclic G_{2n} [illegible]

shall have in the whole group $3n$ cyclic subgroups of order n, whose elements (not contained in G) comprise the remaining half of the $3n^2$ elements GT^jR, $j=1, 2, 3$. But if $n/2$ is even and $n/4$ is odd, there will be $np/2$ elements of period n in GR, p of which will be contained in each G_n. The whole group then will possess $3n/2$ cyclic G_n's, whose elements (exclusive of those in G) comprise one-fourth of the elements GT^jR. The remaining fourth of these elements are distributed equally in $3n$ cyclic groups of order $n/2$. For there are $3np/2$ elements in GT^jR of period $n/2$, and $p/2$ of these occur in each $G_{n/2}$. And so for any even n—we remove the factor 2 repeatedly until the remaining factor is odd. Hence

If n is even and $n/2^a$ is the first odd number in the sequence n, $n/2$, $n/2^2, \cdots$, the $3n^2$ elements GT^jR are distributed as follows:

$3n^2/2$ among $3n/2$ cyclic groups of order $2n$;
$3n^2/2^2$ among $3n/2$ cyclic groups of order n;
$3n^2/2^3$ among $3n/2$ cyclic groups of order $n/2$;
. . . .
$3n^2/2^a$ among $3n/2$ cyclic groups of order $n/2^{a-2}$;
$3n^2/2^a$ among $3n$ cyclic groups of order $n/2^{a-1}$.

Half of the elements of each of these cyclic subgroups (the even powers of the generators) are homologies from G, while the odd powers of the generators are elements GT^jR. Each of the $3n$ cyclic groups of order $n/2^{a-1}$ contains one of the $3n$ reflexions. Note that the number of groups of the several orders is equal except the last which is double that of the others. In particular, if $n=2^a$, the last set of groups are the $3n$ cyclic groups of order two generated by the $3n$ reflexions.

3. *Special Sets.* The special sets of conjugate points and lines under the group are compounded of the fixed points and lines of the individual collineations, presenting themselves in dual configurations. Now every element of G leaves unaltered each vertex and side of the triangle of reference, while each homology of course leaves unaltered every point on one side of the triangle.

The fixed points of the n^2 cyclic G_3's, comprising (besides the identity) the elements GT^j $(j=1, 2)$ aggregate $3n^2$. The axes of reflexion meet by threes at n^2 of these points,—aside from the vertices of the reference triangle. Thus for all values of $i, j=1, \cdots, n$ the point $(\epsilon^i, \epsilon^j, 1)$ lies on the three axes.

$$y-\epsilon^j z=0, \qquad \epsilon^i z-x=0, \qquad \epsilon^j x-\epsilon^i y=0.$$

And obviously only one axis from each vertex can contain any particular point. When $i = j$, which happens n times, the point lies on the axis $x - y = 0$. Evidently the axes are treated symmetrically. Hence, neglecting the intersections at the vertices of the reference triangle,

The $3n$ axes of reflexion meet by threes at n^2 points, n of the points lying on each axis.

Each of the n^2 points is the fixed point of a dihedral subgroup G_6.

For in addition to a cyclic G_3, each is fixed under three reflexions the product of any two of which generate the cyclic G_3. The three centers of these reflexions lie on a line, the fixed line of the dihedral G_6. The n^2 fixed lines of the dihedral groups are dual to the n^2 fixed points while the centers of reflexion are dual to the axes, hence

The $3n$ centers of reflexion lie in sets of n on the sides of the reference triangle and by threes on n^2 lines, n of the lines meeting at each center.

The generator of the cyclic subgroup G_3 of the dihedral G_6 which has the fixed point $(\epsilon^i, \epsilon^j, 1)$ may be written

$$x' = \epsilon^{2i-j}y, \qquad y' = \epsilon^{i+j}z, \qquad z' = x. \tag{6}$$

The two additional fixed points of the G_3 are

$$(\omega\epsilon^i, \omega^2\epsilon^j, 1) \text{ and } (\omega^2\epsilon^i, \omega\epsilon^j, 1), \qquad \omega^3 = 1. \tag{7}$$

If n is not a multiple of 3 the co-ordinates (other than 1) of these points are $3n$-th roots of unity but not n-th roots. And if i, j range over the values of $1, \cdots, n$, we get n^2 cyclic G_3's, each of which has for fixed points one of the n^2 points, and two of a second set of $2n^2$ points. Furthermore, it is obvious from the form of the co-ordinates that both aggregates of points are (special) conjugate sets. If, however, n is a multiple of 3, we get only $n^2/3$ cyclic G_3's by varying i and j, whose fixed points comprise the n^2 points. In this case we have three dihedral G_6's, one associated with each vertex of a triangle of fixed points, and having a common invariant subgroup G_3.

In any case the generators of the n^2 cyclic G_3's may be written

(8) [illegible]

The fixed points and lines of each of the n^2 cyclic G_3's form a self-polar triangle with respect to the invariant triangle of the whole group.

Suppose now that $n \equiv 0$, mod 3. Then if $i + j \equiv 0$, mod 3, (8) represents the $n^2/3$ cyclic G_3's whose fixed points comprise the set of n^2 points. If $i + j \not\equiv 0$, mod 3, (8) represents the $2n^2/3$ cyclic G_3's whose fixed points aggregate $2n^2$. The co-ordinates of the $2n^2$ fixed points as a set may be written

$$(10) \qquad (\eta^{3k+1}, \eta^{3k-1}, 1), \qquad (\eta^{3k-1}, \eta^{3k+1}, 1),$$

where k is an integer, there being n^2 points represented by each expression. All these numbers η^a are $3n$-th roots of unity which are not n-th roots. The co-ordinates of the $2n^2$ points may thus be written in the form

$$(11) \qquad (\epsilon^i\eta, \epsilon^j\eta^2, 1), \qquad (\epsilon^i\eta^2, \epsilon^j\eta, 1), \qquad i, j = 1, \cdots, n,$$

from which it is apparent that they constitute a special conjugate set. We may summarize as follows:

The $3n^2$ fixed points of the n^2 cyclic subgroups G_3 divide into two special conjugate sets of n^2 and $2n^2$ points respectively. If $n \not\equiv 0$, mod 3, the cyclic groups are conjugate, each having for fixed points one of the former and two of the latter. But if $n \equiv 0$, mod 3, the groups themselves divide, one-third and two-thirds into conjugate sets, the fixed points of one being the n^2 points and those of the other the $2n^2$ points. The $3n^2$ fixed lines form dual configurations.

The square of each element GT^jR, $j = 1, 2, 3$, is a homology of G (or the identity). The fixed points of each of the cyclic groups G_r $(r > 2)$ generated by them, accordingly, are a vertex of the triangle of reference and two points on the opposite axis. For example, the fixed points of the group generated by S^iW^kR [equations (4)] which lie on the z-axis are $(\epsilon^{-i}, \pm\epsilon^{k/2}, 0)$. These obviously are cut out by the lines $x^{2n} - y^{2n} = 0$. But half of the points are centers of the n reflexions included among the elements in question. The centers lie on the lines $x^n + y^n = 0$ when n is odd, on $x^n - y^n = 0$ when n is even. Since the sides of the triangle of reference are treated symmetrically, we conclude:

*Each element GT^jR, $j = 1, 2, 3$, leaves fixed one vertex of the invariant triangle of the group. The residual fixed points of the elements make up two special sets of $3n$ conjugate points, lying on the sides of the invariant triangle and cut out by the two special dual sets of conjugate lines,**

* Dual to the sets of points not to each other.

$$(y^n + z^n)(z^n + x^n)(x^n + y^n) = 0, \quad (y^n - z^n)(z^n - x^n)(x^n - y^n) = 0.$$

One set of $3n$ points are centers of reflexion, which are cut out by the first set of lines when n is odd, by the second set of lines (axes) when n is even. This completes the enumeration of special sets of conjugate points—except the obvious sets on the axes of homology and reflexion—which we collect for convenience of reference.

1°. One set of 3, vertices of the invariant triangle;

2°. A set of $3n$, centers of reflexion, cut out of the sides of the invariant triangle by $\Pi(x^n + y^n) = 0$, when n is odd, by $\Pi(x^n - y^n) = 0$, when n is even; *

3°. A counter set of $3n$, cut from the sides of the invariant triangle by $\Pi(x^n - y^n) = 0$, when n is odd, by $\Pi(x^n + y^n) = 0$ when n is even;

4°. One set of n^2, intersections of the axes of reflexion aside from the vertices of the invariant triangle;

5°. One set of $2n^2$, comprising with the n^2 points the fixed points of the n^2 cyclic G_3's;

6°. ∞^1 sets of $6n$, lying on the sides of the invariant triangle, each fixed under a homology of period n;

7°. ∞^1 sets of $3n^2$, lying on the axes of reflexion.

The duals of these are the special sets of conjugate lines, all other points and lines belonging to general conjugate sets.

4. *The Complete System of Invariants.* By the same argument used in the case of the cubic † it is proved that the complete system of invariants of our group consists of the four forms

$$\begin{aligned} s_1 &: \ x^n + y^n + z^n \\ s_2 &: \ y^n z^n + z^n x^n + x^n y^n \\ \Pi_3 &: \ x\,y\,z \\ \Pi_{3n} &: \ (y^n - z^n)(z^n - x^n)(x^n - y^n). \end{aligned}$$

These forms are connected by the syzygy

$$-(\Pi_{3n})^2 \equiv 27s_3^2 + 4s_2^3 + 4s_1^3 s_3 - s_1^2 s_2^2 - 18 s_1 s_2 s_3, \tag{12}$$

where $s_3 = x^n y^n z^n$. Another obvious invariant is

$$\Pi'_{3n}: \ (y^n + z^n)(z^n + x^n)(x^n + y^n) = s_1 s_2 - s_3.$$

* By $\Pi(x^n + y^n)$ we mean the product of three terms of the type within the parentheses.

† Winger, *Tôhoku Mathematical Journal*, Vol. 29 (1928), § 3.

These forms are associated with the covariant system of the binary cubic whose roots are x^n, y^n, z^n:

$$t^3 - s_1 t^2 + s_2 t - s_3. \tag{13}$$

Thus $-(\Pi_{3n})^2$, being the square of the product of differences of the roots, is the discriminant Δ of the cubic,—a fact which leads at once to the syzygy (12).

There is an interesting connection between (a) the covariant system of the cubic, (b) the invariants of the group, and (c) the invariants of the inversion

$$x' = 1/x, \qquad y' = 1/y, \qquad z' = 1/z. \tag{14}$$

The Hessian H and cubicovariant $C_{3,3}$ of the cubic are respectively

$$H: \ (3s_2 - s_1^2)t^2 + (s_1 s_2 - 9s_3)t + (3s_1 s_3 - s_2^2),$$
$$C_{3,3}: \ (-27s_3 + 9s_1 s_2 - 2s_1^3)t^3 + (9s_1 s_3 + s_1^2 s_2 - 6s_2^2)t^2$$
$$+ (-6s_1^2 s_3 + 9s_2 s_3 + s_1 s_2^2)t + (9s_1 s_2 s_3 - 27s_3^2 - 2s_2^3).$$

The binary inversion $t' = 1/t$ has the same effect on the cubic and its covariants as the ternary inversion (14): the coefficients are interchanged in pairs, first with last, second with next to last, etc. Thus *the fundamental forms s_1 and s_2 are interchanged while the systems of lines Π_{3n} and Π_{3n}' and the pencil of curves $s_1 s_2 - \lambda s_3$ are invariant.*

The inversion (14) also leaves invariant as a whole the special set of n^2 points and interchanges in pairs the set of $2n^2$ points (10).

We shall now take up the properties of the fundamental invariants. First we observe that *the curve s_1 has n hyperosculation points on each side of the invariant triangle, whose n tangents meet at the opposite vertex. The points are the centers of reflexion when n is odd and the counter sets of $3n$ when n is even. In both cases the equation of the $3n$ tangents is $\Pi_{3n}' = 0$.*

The hyperosculation points absorb all of the flexes and $\frac{3}{2}n(n-2)(n-3)$ double lines. There remain $\frac{1}{2}n^2(n-2)(n-3)$ double lines. Again

Each of the $3n$ axes of reflexion cuts the curve in n points whose (simple) tangents meet at the corresponding center.

The points and lines are special conjugate sets of $3n^2$. Finally

The $2n^2$ points are the complete intersections of the fundamental invariants s_1 and s_2.

This is proved either by solving simultaneously the equations of the

curves or by substituting the co-ordinates (10) of the points already found. We have found three special sets of points on the curve, all other points occurring in general sets.

We shall now prove two general theorems on curves admitting homologies.

1°. *If a curve is invariant under a homology of period k, the rth polar $(r < k)$ of the center of the homology is composite, containing the axis as a $(k - r)$-fold factor.*

Let the homology be

$$x' = x, \qquad y' = y, \qquad z' = \theta z, \qquad\qquad \theta^k = 1.$$

Then the equation of the curve is of the form

$$f_0 + f_1 z^k + f_2 z^{2k} + f_3 z^{3k} \cdots = 0,$$

where f_i are binary forms in x and y. The rth polar of the center (0, 0, 1) of the homology contains the factor z^{k-r}.

2°. *If a curve of order n admits a homology of period n, the first polar of any point on the axis of homology breaks up into $n - 1$ lines which meet at the center.*

For the multiplier of z in the homology is ϵ, hence the equation of the curve may be written

$$z^n + f^{(n)}(x, y) = 0,$$

where $f^{(n)}$ is a binary form of order n. And the polar of any point $(a, b, 0)$ on the axis of homology is a binary form in x and y of order $n - 1$, which represents a complex of $n - 1$ lines on the center.

Theorem 2°, applied to the curve s_1, yields the theorem:

*The contacts of tangents from an arbitrary point on a side of the invariant triangle fall in sets of n on $n - 1$ lines which meet at the opposite vertex.**

From the inverse relation of s_1 and s_2 we infer at once:

Each vertex of the invariant triangle is for the curve s_2 an r-fold [illegible] *each branch at which the tangent has* [illegible] *point contact. The* [illegible]

and $3n(n-1)(n-2)/2$ double lines in addition to the normal quota. Accordingly the Plücker numbers of the curve are as follows:

Order, $2n$; Class, $n(n+1)$; Genus, $\frac{1}{2}(n-1)(n-2)$; Number of flexes, 0 *; Number of double lines, $\frac{1}{2}n^3(n-1)$.

The n intersections of each axis of reflexion, aside from those at a multiple point, are contacts of simple tangents from the corresponding center, the aggregate of such points forming a special conjugate set of $3n^2$.

The curve s_2 like s_1 has but three special sets of points, including the common set of $2n^2$ points. The $2n^2$ points deserve special attention. We distinguish three cases according as the residue of n with respect to the modulus 3 is 0, 1 or 2.

Suppose first that n is of the form $3k$. Then we saw, § 3, that the $2n^2$ points comprise the fixed points (9) of the $2n^2/3$ cyclic groups (8), $i+j \not\equiv 0$, mod 3. Under the restrictions imposed we find that the three points (9) are the vertices of a Poncelet triangle alike for the curves s_1 and s_2. In fact, denoting the points in the order written by 1, 2, 3, and by (123) indicating that the tangents at the points respectively pass through 2, 3, 1, then the symbol for the triangle on s_1 is (123) when $i+j$ has the form $3k+1$, and (132) when $i+j$ has the form $3k+2$. On s_2 the symbols for the triangle are interchanged under the respective hypotheses. The sides of the triangles are naturally the special sets of $2n^2$ lines on the two curves. Thus

If n is a multiple of 3, the curves s_1 and s_2 have a common system of $2n^2/3$ Poncelet triangles, each the fixed triangle of a cyclic G_3, whose vertices and sides in the aggregate comprise respectively the special sets of $2n^2$ conjugate points and lines.

If n is not a multiple of 3, the fixed points of the n^2 cyclic G_3's are

$$(\epsilon^i, \epsilon^j, 1), \qquad (\omega\epsilon^i, \omega^2\epsilon^j, 1), \qquad (\omega^2\epsilon^i, \omega\epsilon^j, 1) \qquad i, j = 1 \cdot \cdot \cdot n.$$

For a given i, j, the first is an n^2 point while the other two belong to the set of $2n^2$ points. Again, the first is the unique fixed point and the others a conjugate pair of the allied dihedral G_6. The tangents to s_1 at $(\omega\epsilon^i, \omega^2\epsilon^j, 1)$ and $(\omega^2\epsilon^i, \omega\epsilon^j, 1)$ are respectively $\omega^{n-1}\epsilon^{-i}x + \omega^{2n-2}\epsilon^{-j}y + z = 0$ and $\omega^{2n-2}\epsilon^{-i}x + \omega^{n-1}\epsilon^{-j}y + z = 0$. These pass through $(\epsilon^i, \epsilon^j, 1)$ if $n = 3k+2$. But if $n = 3k+1$, they both reduce to

* Except when $n = 2$. The curve is then the projective lemniscate, having three biflecnodes. When n is an even number greater than 2, the branches at the multiple points are sometimes called flexes but properly speaking they are higher singularities.

$$\epsilon^{-i}x + \epsilon^{-j}y + z = 0, \tag{15}$$

which is thus a double line. This situation is just reversed for the curve s_2: (15) is a double line when $n = 3k + 2$; but if $n = 3k + 1$, the tangents at the $2n^2$ points meet in pairs at the n^2 points. Hence, the n^2 cyclic G_3's being conjugate,

If n is of the form $3k + 1$, the special set of n^2 lines (15) *are double lines of s_1 with contacts at the special set of $2n^2$ points* (*cut out by s_2*). *The special set of $2n^2$ lines are tangent to s_2 at the $2n^2$ points and meet in pairs at the n^2 points,—the polars of the n^2 lines with respect to the invariant triangle of the group. When n is of the form $3k + 2$, the rôles of s_1 and s_2 are interchanged.*

We are now in a position to enumerate the double lines which meet at the centers of reflexions, the tangents from a center being either simple, with contacts on the axis, or double lines.

The double lines of s_1. n, odd. Each center is a hyperosculation point, whose tangent is equivalent to n simple tangents. There are also n simple tangents with contacts on the corresponding axis. The remaining tangents combine into $(n^2 - 3n)/2$ double lines. When n is of the form $3k$ or $3k + 2$, these double lines are all distinct. We have then accounted for $\frac{3}{2}n^2(n-3)$ of the double lines, leaving $\frac{1}{2}n^2(n-3)(n-5)$. If n is of the form $3k + 1$, we have just seen that the n^2 lines (15) are double lines. And since each contains three centers of reflexion, each has been counted three times in the above enumeration. Thus while the number of double lines meeting at each center is not affected, the total number involved is reduced to $\frac{1}{2}n^2(3n-13)$, leaving $\frac{1}{2}n^2(n^2 - 8n + 19)$ unaccounted for.

n, even. The centers are not on the curve, so each carries n simple tangents with contacts on the axis, and $\frac{1}{2}n(n-2)$ double lines in addition. Again when n is of the form $3k$ or $3k + 2$, these double lines are all different so that $\frac{1}{2}n^2(n-2)(n-6)$ remain. But if n is of the form $3k + 1$, the n^2 lines are double lines and each is counted three times, leaving $\frac{1}{2}n^2(n-4)^2$ unaccounted for.

We have accounted for all the double tangents of the quartic, quintic and sextic cases. For higher values of n, those which remain fall into general conjugate sets, the number under the appropriate hypotheses in each case, being a multiple of $6n^2$. Finally, in consequence of theorem 2° of this section, we may say

[illegible]

invariant triangle of the group, while the contacts of each set divide into two sets of n lying on lines through the opposite vertices.

The double lines of s_2. *n, odd.* The tangent at each branch of a multiple point passes through a center of reflexion and is equivalent to n simple tangents from the center. From each center also run n simple tangents with contacts on the axis. The remaining tangents unite to form double lines. If n is of the form $3k$ or $3k+1$, there remain $\frac{1}{2}n^3(n-1)-\frac{3}{2}n^2(n-1) = \frac{1}{2}n^2(n-1)(n-3)$. If, however, n is of the form $3k+2$, the n^2 lines are double lines and each is counted three times since it contains three of the centers. In this case the number of residual double lines is $\frac{1}{2}n^2(n^2-4n+7)$.

n, even. The tangents at the multiple points are not on the centers of reflexion. Thus each center carries n simple tangents with contacts on its axis, and in addition $\frac{1}{2}n^2$ double lines. If $n=3k$ or $3k+1$, there remain $\frac{1}{2}n^3(n-4)$. But if $n=3k+2$, on removing duplicates occasioned by the triple counting of the n^2 lines, we are left with $\frac{1}{2}n^2(n-2)^2$.

We have completely accounted for the double lines of s_2 when $n=2, 3, 4$. In all other cases, the residual double lines belong in general conjugate sets, the formula for the number in each case representing a multiple of $6n^2$.

5. *The Construction of Conjugate Sets. The Configuration of a General Set of $6n$ Conjugate Points.* An arbitrary point P of the plane is carried by the homology $x'=x$, $y'=y$, $z'=\epsilon z$, into n points of a line through the center w. Let this line be $\sqrt[n]{a}x+y=0$. The homology $x'=x$, $y'=\epsilon y$, $z'=z$ will carry the line into a set of n lines, $ax^n+y^n=0$, each of which contains n points conjugate to P. These lines in turn are transformed by the reflexion $x'=y$, $y'=x$, $z'=z$, into a second set of n, $x^n+ay^n=0$, each containing n points conjugate to P. Now by a similar process we get two sets of lines from each of the other vertices, each line containing n points conjugate to P. We thus have $6n$ lines, $2n$ on each vertex of the reference triangle, and each carries n points—aggregating $6n^2$—which form a general conjugate set under the group.

Since these lines as a whole are invariant under the group they must be expressible in terms of the complete system. Such a set of $6n$ lines, which we shall call Λ_a, is

$$
\begin{aligned}
(16)\qquad \Lambda_a &\equiv (ay^n+z^n)(y^n+az^n)(az^n+x^n)(z^n+ax^n)(ax^n+y^n)(x^n+ay^n)\\
&\equiv a^3(s_1^2s_2^2-2s_1^3s_3-2s_2^3+4s_1s_2s_3-s_3^2)\\
&\quad +a^2(a^2+1)(s_2^3+3s_3^2+s_1^3s_3-5s_1s_2s_3)\\
&\quad +a(a^2+1)^2(s_1s_2s_3-3s_3^2)+(a^2+1)^3s_3^2=0.
\end{aligned}
$$

For an arbitrary value of the parameter a, the lines of the system are distinct. But when $a = -1$, the lines are the axes of reflexion repeated and Λ_a reduces to $-(\mathrm{II}_{3n})^2$ or Δ. Again when $a = 1$, (16) represents the counter set of $3n$ lines repeated, or $\Lambda_1 - (\mathrm{II}_{3n}')^2$.

We have not yet fully described the configuration of a general set of conjugate points. The n^2 points on the lines $ay^n + z^n = 0$ are invariant as a whole under the two homologies above, hence they lie in sets of n on n lines through each of the vertices v and w of the triangle of reference. Thus we have *a set of n^2 points and $3n$ lines, n lines meeting at each vertex. Each line contains n of the points and each point lies on three of the lines. The points therefore comprise the complete intersections of the lines aside from the vertices of the reference triangle. A general conjugate set consists of six such sets of n^2 points, one set on the lines corresponding to each factor in* (16).

The system of lines Λ_a mutually intersect in $12n^2$ points which in general belong in two special sets of $3n^2$, lying on the system Λ_{-a^2}. Exceptions occur when $a = 1$, or -1 as already noted and also when $a = -\omega$ (or $-\omega^2$), as will appear later.

Λ_a is also an invariant of the inversion (14). Indeed under the inversion each term (function of s_i) in (16) goes into itself. Each, however, can be built up out of the simpler invariants of the inversion, $s_2^3 + s_1^3 s_3 = 0$, and the members of the pencil $s_1 s_2 + \lambda s_3 = 0$.

We shall now consider the problem of the construction of sets of conjugate points on the fundamental invariants s_1 and s_2, utilizing the principle that invariant curves meet in sets of conjugate points.

A second composite invariant is

$$(17) \qquad \begin{aligned} K_c &\equiv (cy^n z^n + x^{2n})(cz^n x^n + y^{2n})(cx^n y^n + z^{2n}) \\ &\equiv cs_2^3 + (c+1)^3 s_3^2 + cs_1 s_3 [cs_1^2 - 3(c+1)s_2] = 0. \end{aligned}$$

Each factor in the parentheses represents a system of double contact conics, belonging to the pencil invariant under a dihedral G_{2n}, whose fixed point-line is a vertex and opposite side of the triangle of reference.

The system of conics K_c cut each other in $6n(3n - 1)$ points, all but $9n^2$ of which coincide at the vertices of the reference triangle. These re-

$$(a^2 - a + 1)^3 s_3^2 + a^2(a-1)^2 s_2^3 = 0 \tag{18}$$

and

$$(c+1)^3 s_3^2 + c s_2^3 = 0. \tag{19}$$

Thence it appears that the system of lines Λ_a as well as the system of conics K_c cut s_1 in the same points as the pencil of invariant curves

$$s_3^2 + \lambda s_2^3 = 0\ * \tag{20}$$

$$\lambda = a^2(a-1)^2/(a^2-a+1)^3 = c/(c+1)^3. \tag{21}$$

The correspondence between the curves of the pencil (20) and the sets of lines Λ_a is one-to-three and not one-to-six as a first glance at relation (21) suggests. For reciprocal values of a determine the same set of lines. Likewise the correspondence between the curves of the pencil and the systems of conics K_c is one-to-three in virtue of (21).

Now observing that when $s_1 = 0$, Δ reduces to $27s_3^2 + 4s_2^3 = 0$, we see that the curve (20) cuts s_1 in the same points as Δ when $\lambda = 4/27$. We have therefore the interesting result:

All sets of conjugate points on s_1, whether special or general, are the complete intersections of the curve with (a) *the systems of lines Λ_a,* (b) *the systems of conics K_c, and* (c) *the pencil of curves* (20).

The special sets are cut out by special curves as follows:

When	the curves Λ_a, K_c, (20) cut out
$\lambda = 0$,	(hyperosculation points) ^{2n};
$\lambda = \infty$,	(the $2n^2$ points)^3;
$\lambda = 4/27$,	(the $3n^2$ points)^2;

where the exponents indicate the multiplicity of intersection. General conjugate sets are cut out by single members of the pencil (20) but by triple sets of lines and conics because of the multiple correspondence. Some exceptions, however, occur in the case of special sets. When $\lambda = \infty$, there is a single set of lines, viz. $a = -\omega$, or $-\omega^2$, cutting out the $2n^2$ points, one line from each vertex of the reference triangle passing through each of the points. When $\lambda = 0$, there are two systems of lines cutting out the hyperosculation points for when $a = 0$ or ∞, the lines reduce to s_3^2, whereas when $a = 1$, the lines become the hyperosculation tangents repeated. Again, when $\lambda = 4/27$, we have from (21) $[(a+1)(a-2)(2a-1)]^2 = 0$, whence there

* It should be noted that this pencil is composite when $n \equiv 0$, mod 3.

are two sets of lines, namely, $a = -1$ and $a = 2$ (or $1/2$). We saw that $\Lambda_1 = \Delta$. We observe now that $\Lambda_2 - \Delta \equiv 9s_1^2 s_2^2$, hence Λ_2 cuts s_1 and s_2 in the same points as Δ. Or, *the system of lines Λ_2 intersect in pairs at the special set of $3n^2$ points both of s_1 and s_2.*

A single set of conics K_0 and K_{-1} respectively cut out the hyperosculation points and the $2n^2$ points. Each conic of the set K_{-1} contains $2n$ of the $2n^2$ points while each of the points is on three of the conics. But when $\lambda = 4/27$, $c = \frac{1}{2}, \frac{1}{2}, -4$ and there are two systems of conics. Two conics of the set $K_{1/2}$ cut s_1 at each of the $3n^2$ points, whereas one conic of the set K_{-4} touches s_1 at each of the points. Furthermore, we find $8K_{1/2} - \Delta \equiv s_1^2(s_2^2 - 2s_1 s_3) \equiv s_1^2(y^{2n}z^{2n} + z^{2n}x^{2n} + x^{2n}y^{2n})$, hence the remaining intersections of the system $K_{1/2}$ with the axes of reflexion fall on a curve of the type s_2 of double the order.

A second solution of the problem, involving the special sets symmetrically, is not without interest. We may regard each special set of points as composed of $6n^2$, each point counted 2, 3 or $2n$ times. Then recalling that these special sets of points are cut out by Δ, s_2^3 and s_3^2, and taking J for a parameter, we have the result:

All sets of conjugate points on s_1 are determined by the system of curves

$$J : J - 1 : 1 = -27s_3^2 : \Delta : 4s_2^3.* \tag{22}$$

The construction of conjugate sets on s_2 is entirely analogous. We should naturally seek a pencil of curves of order $3n$. However, the systems of lines Λ_a and conics K_c are still available, for half of the intersections with s_2 fall at the multiple points, leaving in general $6n^2$ free intersections. We shall state the results at once:

All sets of conjugate points on s_2, both special and general, are cut out by (a) *the systems of lines Λ_a,* (b) *the systems of conics K_c,* (c) *the pencil of curves*

$$s_3 + \lambda s_1^3 = 0,\dagger \tag{23}$$

the parameters satisfying the relations

$$\lambda = a^2(a-1)^2/(a^2 - a + 1)^3 = c^2/(c+1)^3. \tag{24}$$

When	the several curves cut out
$\lambda = 0$,	(the three multiple points)$^{2n^2}$;
$\lambda = \infty$,	(the $2n^2$ points)^3;
$\lambda = 4/27$,	(the $3n^3$ points)^2.

* Cf. Dyck, *loc. cit.*, p. 51[illegible].

† The pencil is composite when n is a multiple of 3.

The correspondence between the curves of the pencil (23) on the one hand and the systems of lines and conics on the other is one-to-three in general but exceptions occur in the case of special sets. When $\lambda = 4/27$, there are but two systems of conics for then $c = 2, 2, -1/4$. The system K_2 pass two and two through the $3n^2$ points, each conic containing $2n$ of the points. The remaining intersections of these conics with axes of reflexion fall on a curve like s_1 but of twice the order, for

$$K_2 - \Delta \equiv s_2^2(s_1^2 - 2s_2) \equiv s_2^2(x^{2n} + y^{2n} + z^{2n}).$$

Each conic of the set $K_{-1/4}$ touches s_2 at n of the $3n^2$ points.

Again, *all sets of conjugate points on s_2 are determined by the pencils*

$$(25) \qquad J : J - 1 : 1 = -27s_3^2 : \Delta : 4s_1^3 s_3,$$

where the points cut out by s_3 appear as extraneous in each set, the pencils being of order $6n$.

The inverse pencils of invariant curves $s_1^2 + \lambda s_2 = 0$ and $s_2^2 + \lambda s_1 s_3 = 0$ contain numerous members with intersecting geometrical properties as exemplified by the rather extensive discussion for the case $n = 3$.*

6. *Extension to Higher Spaces.* Many of the properties of the group and its invariants are susceptible of immediate generalization. Thus in space S_r of r dimensions we should be concerned with the monomial group that leaves invariant the variety

$$(26) \qquad x_0^n + x_1^n + x_2^n + \cdots + x_r^n = 0.$$

The order of the group is $(r+1)!\,n^r$, comprising the elements

$$(27) \quad x_0' = x_0, \quad x_1' = \epsilon^{i_1} x_1, \cdots, x_r' = \epsilon^{i_r} x_r, \qquad i_1, i_2, \cdots, i_r = 1, \cdots, n,$$

combined with those of the group of order $(r+1)!$ which permutes the variables in all possible ways. The group is thus in $(n^r, 1)$ isomorphism with the permutation group $G_{(r+1)}!$ on $r+1$ letters,—a fact which supplies a key to its essential structure. The n^r elements (27) constitute an invariant Abelian subgroup which corresponds to the identity in the multiple isomorphism.

The associated binary form is now of order $r+1$:

$$(28) \qquad t^{r+1} - s_1 t^r + s_2 t^{r-1} - \cdots + (-1)^r s_r t + (-1)^{r+1} s_{r+1} = 0,$$

* Winger, *loc. cit.*, § 6. There is one curve of each system on the n^2 points, viz. $\lambda = -3$. This curve in each pencil is composite:

$$s_1^2 - 3s_2 \equiv (wx^n + w^2y^n + z^n)(w^2x^n + wy^n + z^n)$$

and the other is the inverse of this.

whose roots are $x_0^n, x_1^n, \cdots, x_r^n$. The complete system of invariants of the group comprises the r elementary symmetric functions of the roots of (28), $s_1, s_2, \cdots, s_r$, together with the product of the variables $x_0 x_1 \cdots x_r$,* and the difference product

$$\begin{aligned}\Pi \equiv (x_0^n - x_1^n)(x_0^n - x_2^n)(x_0^n - x_3^n) \cdots (x_0^n - x_r^n) \\ (x_1^n - x_2^n)(x_1^n - x_3^n) \cdots (x_1^n - x_r^n) \\ (x_2^n - x_3^n) \cdots (x_2^n - x_r^n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ (x^n_{r-1} - x_r^n).\end{aligned}$$

The square of Π is the discriminant of the binary form, an identity which implies a syzygy among the fundamental forms. Every invariant as well as each coefficient of every covariant of the binary form is an invariant of the group. Indeed from each such invariant of the group can be derived a family of invariants by inserting a parameter in every term.

The inversion

$$(29) \qquad x_i' = 1/x_i, \qquad i = 0, 1, \cdots, r$$

has the same effect on the form (28) as the binary inversion $t' = 1/t$. Thus the pairs of invariants s_i, s_{r-i+1} are interchanged. Accordingly we may construct invariants of the inversion by forming isobaric expressions such as

$$(30) \qquad s_i s_{r-i+1} + \lambda s_{i+1} s_{r-i},$$

where λ is a parameter. Such invariants of the inversion can also be constructed from the coefficients of every invariant of the binary form. In particular, if any covariant of the binary form (including the form itself) is of even order, the middle coefficient is an invariant of the inversion. Again, every invariant of the binary form is an invariant of the inversion. In fact, Π, the square root of the discriminant of the binary form, is an invariant of the inversion. Thus the interrelation between the covariant system of the binary form and the invariants of the group and the inversion, noted in the plane, carries on to higher dimensions.

The collineation group in S_r effects on each of the $r+1$ reference spaces S_{r-1} a collineation group of the type in question for the next lower space [illegible]

In S_3 the group is in $(n^3, 1)$ isomorphism with the octahedral G_{24}.* The elements of the Abelian invariant subgroup are given by (27) when $r = 3$. The binary form (28) now reduces to the quartic

$$t^4 - s_1 t^3 + s_2 t^2 - s_3 t + s_4. \tag{31}$$

The fundamental invariants are the surfaces

$$\begin{aligned}
\Pi_4:&\quad x_0 x_1 x_2 x_3\\
s_1:&\quad x_0^n + x_1^n + x_2^n + x_3^n\\
s_2:&\quad x_0^n x_1^n + x_0^n x_2^n + x_0^n x_3^n + x_1^n x_2^n + x_1^n x_3^n + x_2^n x_3^n\\
s_3:&\quad x_0^n x_1^n x_2^n + x_0^n x_1^n x_3^n + x_0^n x_2^n x_3^n + x_1^n x_2^n x_3^n\\
\Pi_{6n}:&\quad (x_0^n - x_1^n)(x_0^n - x_2^n)(x_0^n - x_3^n)(x_1^n - x_2^n)(x_1^n - x_3^n)(x_2^n - x_3^n),
\end{aligned}$$

of which obviously Π_4 is the tetrahedron of reference, while Π_{6n} represents a system of $6n$ planes, passing in sets of n through the edges of the tetrahedron.

The inversion $x_i' = 1/x_i$, $i = 0, 1, 2, 3$ interchanges the surfaces s_1 and s_3 but leaves s_2 unaltered. Other invariants of the inversion are: the system of planes Π_{6n}, the two invariants of the binary quartic, the middle term in the Hessian and the sextic covariant of the quartic. The intersections of s_2 with the surfaces s_1 and s_3 are a pair of space curves invariant under the group but interchanged by the inversion. The intersection of s_1 and s_3 however is a curve invariant alike under the group and the inversion. The reference planes cut the surface s_1 and s_2 respectively in sets of four curves of the types $x^n + y^n + z^n$ and $y^n z^n + z^n x^n + x^n y^n$. Each curve, of course, is invariant under the ternary group in its plane, while the four curves in either set are a conjugate set under the quaternary group.

In this section we have sketched in very rough outline the problem for the higher spaces. A systematic investigation of the group and its invariants for an arbitrary value of n even in S_3 would doubtless yield much of interest alike for geometry and group theory.

UNIVERSITY OF WASHINGTON.

* For a treatment of this group when $n = 6$, see Musselman, *loc. cit.*

Plane Involutions of Order Three or Four.

By F. R. Sharpe.

1. In "The problem of Plane Involutions of Order $t > 2$," * the following general method of constructing families of involutions was given. Let $|C| = a_1C_1 + a_2C_2 = 0$ be a pencil of curves of genus p. Construct a system of curves $|C| + |\bar{C}|$ which contains $|C|$ with a residual $|\bar{C}|$ and such that a curve C_3 of $|C| + |\bar{C}|$ meets a curve of $|C|$ in t variable points. The equations $y_1 = C_1\bar{C}$, $y_2 = C_2\bar{C}$, $y_3 = C_3$ define an involution of order t from which a family of involutions can be derived if the dimensions R of $|C| + |\bar{C}| > 2$. The cases in which $R = 2$ were not considered. It will be shown that they lead to new involutions, particularly if $t = 3$ or 4.

2. If two systems $|C|$ and $|\bar{C}|$ having i variable intersections, are of grade (number of variable intersections) $n, \bar{n}$, genus $p, \bar{p}$ and vitual dimension $\gamma, \bar{\gamma}$ then for $|C| + |\bar{C}|$, $N = n + \bar{n} + 2i$, $P = p + \bar{p} + i - 1$, $R = r + \bar{r} + i$.

The pencil $|C|$ is of vitual dimension $-p + 1 + n$ and grade n if $n \leq p$ of its simple basis points are not basis points of $|C| + |\bar{C}|$ and $i = t - n$. The dimension of $|C|$ must be 0. The dimension of $|C| + |\bar{C}|$ is therefore $2 = (-p + 1 + n) + 0 + (t - n)$. Hence $p = t - 1$. Moreover $P = p + \bar{p} + t - n - 1$, and $N = t$ when the y points in which C_3 meets C are included in the basis points. For a regular system $R = N - P + 1$ and the super abundance $s = R - (N - P + 1)$. Hence $s = y = P + 1 - t$ $= p + \bar{p} - n$.

3. Let $t = 3$ so that $p = 2$. With the pencil $C_4 : A^2 12B$ we find

$\|C\|$	$\|\bar{C}\|$	$\|C\| + \|\bar{C}\|$
$C_4 : A^2 8B3C$	$C_3 : A8B$	$C_7 : A^3 8B^2 3C2D$
$C_4 : A^2 9B3C$	$C_3 : 9B$.	$C_7 : A^2 9B^2 3C3D$
$C_4 : A^2 B10C$	$C_4 : AB^2 10C$	$C_8 : A^3 B^3 10C^2 3D$
$C_4 : A^2 7B3C2D$	$C_6 : A^2 7B^2 3C$	$C_{10} : A^4 7B^3 3C^2 2D4E$
$C_4 : A^2 9B2C$	$C_7 : A^3 9B^2 2C$	$C_{11} : A^5 9B^3 2C^2 4D$
$C_4 : A^2 B8C2DE$	$C_7 : A^2 B^3 8C^2 2D$	$C_{11} : A^4 B^4 8C^3 2D^2 E5F$
$C_4 : A^2 11BC$	$C_7 : A11B^2 C$	$C_{11} : A^3 11B^3 C^2 6D$
$C_4 : A^2 8B2CDE$	$C_{10} : A^4 8B^3 2C^2 D$	$C_{14} : A^6 8B^4 2C^3 D^2 E6F$.

* F. R. Sharpe, *American Journal of Mathematics*, Vol. 50 (1928), pp. 627-635.

Thus in the first case $n=1$, $p=1$, $P=4$, $y=2$ and similarly for the others.

With the pencil $C6: 8A^2BB'CC'$ we find

$\mid C \mid$	$\mid \overline{C} \mid$	$\mid C \mid + \mid \overline{C} \mid$
$C6: 2A^2 6B^2 CC'D$	$C_1: 2A$	$C_7: 2A^3 6B^2 CC'DE$
$C6: A^2 7B^2 CC'DD'$	$C_1: AC$	$C_7: A^3 7B^2 C^2 C'DD'2E$
$C6: A^2 7B^2 CC'D$	$C_3: 7BCD$	$C_9: A^2 7B^3 C^2 D^2 C'2E$
$C6: 8A^2 CC'DD'$	$C_4: 8AC^2 C'DD'$	$C_{10}: 8A^3 C^3 C'^2 D^2 D'^2 4E$

With the pencil $C_7: A^3 10B^2$

$C_7: A^3 2B^2 8C^2$	$C_1: 2B$	$C_8: A^3 2B^3 8C^2 2D$
$C_7: A^3 9B^2 C^2$	$C_3: 9B$	$C_{10}: A^3 9B^3 C^2 3D$

With the pencil $C_9: 8A^3 2B^2 C$

$C_9: 2A^3 6B^3 2C^2 D$	$C_1: 2A$	$C_{10}: 2A^4 6B^3 2C^2 D2F$
$C_9: A^3 7B^3 C^2 D^2 E$	$C_3: 7BCE$	$C_{12}: A^3 7B^4 C^3 D^2 E^2 3F$
$C_9: A^3 7B^3 2C^2$	$C_3: 7B2C$	$C_{12}: A^3 7B^4 2C^3 2D$
$C_9: 6A^3 2B^3 2C^2 D$	$C_6: 6A^2 2B2C^2 D$	$C_{15}: 6A^5 2B^4 2C^4 D^2 4E$

With $C_{13}: A^5 9B^4$

$C_{13}: A^5 9B^4$	$C_3: 9B$	$C_{16}: A^5 9B^5 3C$

If $t=4$, $p=3$. With the pencil $C_4: 16A$

$\mid C \mid$	$\mid \overline{C} \mid$	$\mid C \mid + \mid \overline{C} \mid$
$C_4: A11B3C$	$C_4: A^2 11B$	$C_8: A^3 11B^2 3C4D$
$C_4: A14B$	$C_5: A^3 14B$	$C_9: A^4 14B^2 5C$
$C_4: AB14C$	$C_6: A^4 B^2 14C$	$C_{10}: A^5 B^3 14C^2 6D$
$C_4: 7A6B3C$	$C_6: 7A^2 6B$	$C_{10}: 7A^3 6B^2 3C6D$
$C_4: 6A9B$	$C_6: 6A^2 9B$	$C_{10}: 6A^3 9B^2 6C$
$C_4: 11A2B3C$	$C_7: 11A^2 2B$	$C_{11}: 11A^2 2B^2 3C7D$
$C_4: A8B5C2D$	$C_7: A^3 8B^2 5C$	$C_{11}: A^4 8B^3 5C^2 2D7E$
$C_4: 10A5B$	$C_7: 10A^2 5B$	$C_{11}: 10A^3 5B^2 7C$

and similarly for the other pencils of genus 3.

CORNELL UNIVERSITY.

A Special Prepared System for two Quadratics in N Variables.

By J. Williamson.

Introduction. A complete system of concomitants for two quadratics in n variables is a system of concomitants, in terms of which every rational integral concomitant of the two quadratics may be expressed rationally and integrally. The existence of such a finite complete system is a particular case of Gordan's general theorem.* There are three possible types of complete systems,

$$
\begin{aligned}
G &= G(a, r, u_1, u_2, \cdots, u_{n-2}, x),\\
H &= H(a, r, x_1, x_2, \cdots, x_{n-1}, x_n),\\
K &= K(a, r, \pi_1, \pi_2, \cdots, \pi_{n-1}, x),
\end{aligned}
$$

where each contains coefficients a, r of the two quadratic ground forms, the u_i denote plane co-ordinates, x and x_i denote point co-ordinates, and π_i denote compound co-ordinates. From a geometrical point of view the K system is the most important, but also it is the largest and the most difficult to determine. This K system is known for the cases $n = 2$,† $n = 3$,‡ and $n = 4$.§ The number of irreducible concomitants for the cases $n = 2$, $n = 3$, and $n = 4$ are 6, 20 and 122 respectively. For the cases $n = 2$ and $n = 3$, the system is strictly irreducible and thus the complete system in these two cases is also the minimum system.¶

This system K is not the most general system, since such a system would contain other sets of variables ρ_i, σ_i, $\cdots$, all cogredient to π_i. But by Clebsch's theorem ‖ all concomitants involving such sets of variables ρ_i, σ_i, etc. can be deduced from the K system by polarization, if to the system K is added the actual concomitant of the field, the determinant of X (see § 1).

---

* Grace and Young, *Algebra of Invariants*, Chap. 6.

† Grace and Young, *loc. cit.*, page 161.

‡ Grace and Young, *loc. cit.*, pp. 280-286.

§ H. W. Turnbull, "The Simultaneous System of Two Quadratic Quaternary [illegible] Quadratics," [illegible]

[illegible] quadratische Formen." [illegible]

The G system has been determined * and in a joint paper by H. W. Turnbull and the author,† a complete system is found, when the co-ordinates π_i (or p_i) are decomposed into their components u or x.

In this paper we are interested in the K system and, though no complete system is determined in the general case, a distinct step is made in that direction. If every concomitant of the two quadratics may be expressed as a product of symbolic factors, where each symbol occurs an even number of times in each product and distinguishing marks on equivalent symbols may be neglected, the totality of such factors is said to form a prepared system for the two quadratics. We determine here a prepared system, in terms of which every concomitant, multiplied by a suitable invariant factor, may be expressed. This system is comparatively simple and consists of $2^n - 1$ factors. In obtaining a prepared system, in terms of which every concomitant, without being multiplied by an invariant factor, may be expressed new factors must be introduced, when $n \geqq 4$. The number of such new factors is 1, 8, and 52, for the cases $n = 4$, 5, and 6 respectively. It is hoped that these results will be published later.

1. *Co-ordinate Systems.* Let $(u) = (u^1 u^2 \cdots u^n)$ denote a set of n independent variables and suppose that (u) represents hyperplane co-ordinates in a space of $n - 1$ dimensions. If $(u_1) = (u)$, $(u_2), \cdots, (u_n)$ are n sets of cogredient variables, such that the sets (u_i) are linearly independent, we have an n-rowed square matrix

$$U = \| u_i^j \|$$

such that the determinant of U is different from zero; i. e.

$$D = | U | \neq 0. \tag{1}$$

A complete co-ordinate system for this space of $n - 1$ dimensions is given by

$$\pi_r = (u_1 u_2 \cdots u_r) \qquad (r = 1, 2, \cdots, n-1),$$

where π_r denotes the set of $n!/r!(n-r)!$ determinants formed from the first r columns of U. If

$$X = \| x_i^j \|$$

is the n-rowed square matrix, whose elements x_i^j are the co-factors of the elements u_{n+1-i}^{n+1-j} in D, then by Jacobi's Ratio Theorem

* H. W. Turnbull, "The Irreducible Concomitants of Two Quadratics in n Variables," *The Transactions of the Cambridge Philosophical Society*, Vol. 21, No. 8 (July, 1909), pp. 197-240.

† H. W. Turnbull and J. Williamson, "Further Invariant Theory of Two Quadratics in n Variables," *Proceedings of the Royal Society of Edinburgh*, Vol. 50, Part I, No. 2 (1929-30), pp. 8-25.

$$\pi_r = (u_1 u_2 \cdots u_r) = D^{r+1-n}(x_1 x_2 \cdots x_{n-r}) = D^{r+1-n} p_{n-r},$$

where p_{n-r} is the set of $n!/r!(n-r)!$ determinants formed form the first $n-r$ columns of X. Hence we have the two dual co-ordinate systems: *

$$\begin{aligned}
(u) &= \pi_1 = D^{2-n} p_{n-1}, \\
(u_1 u_2) &= \pi_2 = D^{3-n} p_{n-2}, \\
&\cdots\cdots\cdots \\
(u_1 u_2 \cdots u_{n-2}) &= \pi_{n-2} = D^{-1} p_2, \\
(u_1 u_2 \cdots u_{n-1}) &= \pi_{n-1} = (x_1) = (x),
\end{aligned}$$

and finally

$$|X| = D^{n-1}.$$

2. *Notation.* Let

(2) $$f = a_x^2 = b_x^2 = c_x^2 = \cdots, \quad g = r_x^2 = s_x^2 = t_x^2 = \cdots$$

be the symbolic forms of two quadratics in n variables, where $a_x = \sum_{i=1}^{n} a_i x_i$. Further let $A_i = B_i = C_i$ denote a matrix of n rows and i columns formed by i equivalent symbols of f, and $R_i = S_i = T_i$ a matrix of n rows and i columns formed by i equivalent symbols of g. We then say that A_i or R_i is of currency i. If, also, X_k denotes a matrix of n rows and k columns formed by k sets of variables cogredient with (x), and U_k a similar matrix formed by k sets of cogredient variables, each contragredient to the set (x), we may denote the ordinary determinantal bracket factor by

$$(A_i R_j U_k) \qquad (i + j + k = n,\ i \geqq 0,\ j \geqq 0,\ k \geqq 0).$$

We shall however require other types of bracket factors: I. Compound Inner Products; II. Generalised Outer Products; III. Generalised Compound Inner Products.

I. *Compound Inner Products.* The Compound inner product

(3) $$(A_i R_j \mid X_{i+j}) \qquad (i + j \leqq n,\ i \geqq 0,\ j \geqq 0)$$

is a determinant of $i + j$ rows and columns and is equal to

(4) $$|a_x b_y \cdots r_z s_w \cdots| = \Sigma \pm a_x b_y \cdots r_z s_w \cdots,$$

where $A_i = ab \cdots$, $R_j = rs \cdots$, $X_{i+j} = xy \cdots zw \cdots$. For example

$$(ab \mid xy) = \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} = a_x b_y - b_x a_y.$$

* For a more detailed account of dual co-ordinate systems, see Turnbull, "Determinants, Matrices and Invariants," pp. 86-88.

The determinantal permutation indicated by $\sum \pm$ in (4) is often written

$$\dot{a}_x \dot{b}_y \cdots \dot{r}_z \dot{s}_w \cdots \text{ or } (a \mid \dot{x})(b \mid \dot{y}) \cdots (r \mid \dot{z})(s \mid \dot{w}) \cdots,$$

where the dots mean that the letters beneath them must be permuted in every possible way, a change in sign being made with each transposition. If $i + j = n$ in (3),

$$(A_i R_j \mid X_n) = (A_i R_j) \mid X_n \mid,$$

and if $i + j = 1$, this inner product reduces to the ordinary simple inner product $(a \mid x) = a_x$ or $(b \mid x) = b_x$. Since the inner product (3) is a determinant, it may be expanded by Laplacian developments in various ways. In fact,

$$(A_i R_j \mid X_i Y_j) = (\dot{A}_i \mid X_i)(\dot{R}_j \mid Y_j) = (A_i \mid \dot{X}_i)(R_j \mid \dot{Y}_j),$$

where again the dots indicate a series of $(i + j)!/i!j!$ terms obtained by interchanging the symbols of A_i with those of R_j in every possible way (or else those of X_i with those of Y_j). Once again each interchange is accompanied by a change of sign. For example,

$$\begin{aligned}(abc \mid xyz) &= (a\dot{b} \mid xy)(\dot{c} \mid z) = (ab \mid \dot{x}\dot{y})(c \mid \dot{z}),\\ &= (ab \mid xy)(c \mid z) - (ac \mid xy)(b \mid z) - (cb \mid xy)(a \mid z),\\ &= (ab \mid xy)(c \mid z) - (ab \mid xz)(c \mid y) - (ab \mid zy)(c \mid x).\end{aligned}$$

It should be noted, that the dot notation does not mean that the symbols appearing in the same bracket factor should be interchanged.* Moreover

(5) $$(A_i R_j \mid p_{i+j}) = D^{n-i-j-1}(A_i R_j \pi_{n-i-j}).$$

II. *Generalised Outer Products.* The generalised outer product

(6) $$(A_i R_j \mid X_k) \qquad (i + j \geqq n,\ k = i + j - n)$$

is defined as $(\dot{B}_{n-j} R_j)(\dot{C}_k \mid X_k)$ where $B_{n-j} C_k = A_i$ and the dots indicate a series of $i!/k!(n - j)!$ terms. By the elementary fundamental identities,

$$(\dot{B}_{n-j} R_j)(\dot{C}_k \mid X_k) = (A_i \dot{S}_{n-i})(\dot{T}_k X_k), \text{ where } T_k S_{n-i} = R_j.$$

The simplest type of such a factor is

(7) $$(A_i R_{n+1-i} x) = (\dot{A}_{i-1} R_{n+1-i}) \dot{a}_x \qquad (i = 2, 3, \cdots, n - 1),$$

where $A_i = A_{i-1} a$.

III. *Generalised Compound Inner Products.*

---

* Turnbull, *loc. cit.*, pp. 83 sq.

If for brevity in (7) we denote $(A_i R_{n+1-i} x)$ symbolically by i_x, we can form a third type of factor, a generalised compound inner product, which is defined in a manner analogous to the definition of a compound inner product. The generalised compound inner product $(ij \ldots k \mid xy \cdots z)$ is a determinant of m rows and columns, where m is the number of symbols $i, j, \cdots, k$. It is equal to

$$| i_x j_y \cdots k_z | = \dot{i}_x \dot{j}_y \cdots \dot{k}_z = (i \mid \dot{x})(j \mid \dot{y}) \cdots (k \mid \dot{z}).$$

3. *Determination of a Prepared System.* By the fundamental theorem of invariants, every concomitant T of the two quadratics (2) can be represented as a sum of terms, where each term is a product of factors of the types $(d_1 d_2 \cdots d_n)$, $(d_1 d_2 \cdots d_i \pi_{n-i})$, and d_x, together with D, and each symbol d represents a symbol of f or g and must occur exactly twice in every product. But by (5)

$$(8) \qquad D^{n-1}(d_1 d_2 \cdots d_n) = | X | \, (d_1 d_2 \cdots d_n) = (d_1 d_2 \cdots d_n \mid X),$$

and

$$(9) \qquad D^{n-i-1}(d_1 d_2 \cdots d_i \pi_{n-i}) = (d_1 d_2 \cdots d_i \mid p_i).$$

Hence if M be any such term of T,

$$(10) \qquad D^k M = K,$$

where K is a product of factors of the types on the right of equations (8) and (9) and k is the total number of variables x_i distinct from x, which occur in M. We now prove:

LEMMA I. *If K has the equivalent symbols $a, b, \cdots, c$, i in number, convolved together as A_i in the same factor, the complementary symbols $a, b, \cdots, c$ appearing in K may also be convolved together as A_i.*

Since every factor occurring in K is a compound inner product, we may write,

$$K = \Sigma (A_i \mid X_i) a_x b_y \cdots c_z N.$$

As $a, b, \cdots, c$ are equivalent symbols, by permuting them determinantally we get

$$i! \, K = \Sigma (A_i \mid X_i)(A_i \mid Y_i) N, \text{ where } Y_i = (xy \cdots z).$$

This proves the lemma. We apply the lemma successively for the cases $i = n$, $n - 1$, etc. and finally have the result:

LEMMA II. *The concomitant K is zero, or else it can be expressed as a sum of terms, where each term is a product of factors of the two types*

$(A_i \mid Y_i)$, $(R_j \mid Y_j)$ *and each symbol* A_i, R_j *occurs exactly twice in each product.*

Further, if K has the factor $(A_n \mid X_n)$ or $(R_n \mid X_n)$, K has the invariant factor $(A_n \mid X_n)^2$ or $(R_n \mid X_n)^2$, or else is zero, since there are only n sets of variables (x). In either case we say that K is reducible.

Now two quadratics in n variables have the $n+1$ invariants

$$(11) \qquad (A_i R_{n-i})^2 \qquad (i = 0, 1, \cdots, n),$$

which form a complete system of invariants,* and n quadratic covariants: the two original quadratics a_x^2, r_x^2 and the $n-2$

$$(12) \qquad (A_i R_{n+1-i} x)^2 \qquad (i = 2, 3, \cdots, n-1)$$

defined by (7).† Let us denote the quadratic covariants symbolically by 1_x^2 and n_x^2 for a_x^2 and r_x^2 respectively, and by i_x^2 for $(A_i R_{n+1-i} x)^2$ $(i = 2, 3, \cdots, n-1)$. We require the following lemma:

LEMMA III. *The generalised inner product*

$$(i, i+1, i+2, \cdots, i+k \mid X_{k+1}) = \prod_{j=i}^{i+k-1} (A_j R_{n-j}) (A_{i+k} R_{n+1-i} \mid X_{k+1})$$

together with additional terms.

For, if $s \geqq 0$ and $A_{i+k} = A_k A_i$,‡

$$(\dot{A}_i R_{n-i}) (\dot{A}_k B_{i-s} U_{n-i+s-k})$$

$$= (B_{i-s} \dot{U}_s R_{n-i}) (A_{i+k} \dot{U}_{n-i-k}) + \sum_{t=1}^{i-s} (\dot{B}_{i-s-t} \overline{U}_{s+t} R_{n-i}) (A_{i+k} \dot{B}_t \overline{U}_{n-i-k-t}),$$

$$= (B_{i-s} \dot{U}_s R_{n-i}) (A_{i+k} \dot{U}_{n-i-k}) + \sum_{t=1}^{i-s} (B_{i-s-t} \dot{U}_{s+t} R_{n-i}) (A_{i+k+t} \dot{U}_{n-i-k-t}),$$

$$= (B_{i-s} \dot{A}_s R_{n-i}) (\dot{A}_{i+k-s} U_{n-i-k-s}) + \sum_{t=1}^{i-s} (B_{i-s-t} \dot{A}_{s+t} R_{n-i}) (\dot{A}_{i+k-s} U_{n-i-k+s}),$$

where $A_{i+k} B_t = A_{i+k+t}$. In all but the first term the currency of A_i has been increased at the expense of the currency of B_{i-s}, which is less than or equal to that of A_i. By (5) we may change the variables throughout and write

$$(13) \qquad (\dot{A}_i R_{n-i}) (\dot{A}_k B_{i-s} \mid X_{i+k-s}) \equiv (B_{i-s} \dot{A}_s R_{n-i}) (\dot{A}_{i-k-s} \mid X_{i+k-s})$$

---

* Turnbull, *loc. cit.*, page 304.

† Turnbull and Williamson, "The Minimum System of Two Quadratics in n Variables," *Proceedings of the Royal Society of Edinburgh*, Vol. 45, Part 2 (1924-25), pp. 149-165.

‡ To avoid confusion a bar is sometimes used in place of a dot to denote determinantal permutations.

mod increased currency. We use $\equiv$ to denote "equal to except for terms, which may be neglected," and the words following *mod* to indicate why these terms may be neglected. We proceed to prove lemma III by induction, assuming its truth for the value k. Hence, if

$$L=(i, i+1, \cdots, i+k, i+k+1 \mid X_{k+2}),$$

by hypothesis

$$L \equiv \prod_{j=i}^{i+k-1}(A_j R_{n-j})(\dot{A}_{i-1} R_{n+1-i})(\bar{B}_{i+k} R_{n-i-k})(\dot{A}_{k+1} \bar{b} \mid X_{k+2}),$$

$$\equiv \prod_{j=i}^{i+k-1}(A_j R_{n-j})(\dot{A}_{i-1} R_{n+1-i})(\bar{B}_{i-1} \dot{A}_{k+1} R_{n-i-k})(\bar{B}_{k+2} \mid X_{k+2}) \text{ by (13)},$$

$$\equiv \prod_{j=i}^{i+k-1}(A_j R_{n-j})(\bar{B}_{i-1} R_{n+1-i})(A_{i+k} R_{n-i-k})(\bar{B}_{k+2} \mid X_{k+2})$$

mod increased currency of R_{n+1-i} by transferring A_{i-1} from the second to the third factor. Accordingly

$$L \equiv \prod_{j=i}^{i+k}(A_j R_{n-j})(\dot{B}_{i-1} R_{n+1-i})(\dot{B}_{k+2} \mid X_{k+2}).$$

In addition

$$(i, i+1 \mid xy)=(\dot{A}_{i-1} R_{n-i+1})(\bar{B}_i S_{n-i})(\dot{a} \bar{b} \mid xy),$$

$$=(A_i S_{n-i})(\bar{B}_{i-1} R_{n-i+1})(\bar{B}_2 \mid xy)$$

by putting $k=1$ in the previous work. Hence the proof by induction is complete except when $i=1$ or $i+k=n$. But, if $i=1$, the term $(A_{i-1} R_{n+1-i})$ disappears and by an otherwise similar proof

$$(14) \qquad (1, 2, \cdots, k \mid X_k) \equiv \prod_{i=1}^{k-1}(A_i R_{n-i})(A_k \mid X_k),$$

and similarily

$$(15) \qquad (n, n-1, \cdots, n-k+1 \mid X_k)=\prod_{i=1}^{k-1}(R_i A_{n-i})(R_k \mid X_k).$$

Though not necessary at this stage, it is important to notice for future applications that the terms neglected in lemma III have all had their currency raised.

It follows from lemma III that, if K be multiplied by the invariant factor $\prod_{i=1}^{k-1}(A_i R_{n-i})^2$ for every factor $(A_k \mid X_k)(A_k \mid Y_k)$ occurring in it and by $\prod_{i=1}^{k-1}(R_i A_{n-i})^2$ for every factor $(R_k \mid X_k)(R_k \mid Y_k)$ occurring in it, and if I denote this total invariant factor, IK can be expressed as a sum of terms, where each term is a product of factors of the types $(1, 2, \cdots, k \mid X_k)$ and [illegible]

symbol $a, b, \cdots, r, s, \cdots$ occurring in K by a symbol i occurring in IK. We can therefore convolve the variables back again exactly as they were originally convolved in K and have the final result:

The concomitant $IK = \Sigma N$, *where* N *is a product of factors of the type* $(i, j, \cdots, k \mid p_m)$ *and each symbol* $i, j, \cdots, k$ *must occur exactly twice in each product.*

But each $N = \Sigma I_i G_i$, where I_i is an invariant factor composed of powers of $D_j = (A_j R_{n-j})^2$, and G_i is a concomitant involving the symbols A_i, R_j convolved in pairs. Hence, if we determine the complete system for the G_i, we have determined a system in terms of which every concomitant of the two quadratics, if multiplied by an invariant factor composed of powers of D_i $(i = 1, 2, \cdots, n-1)$, can be expressed.

Order in which the concomitants G_i *are considered.* Let $\mu_{i,j}$ denote the number of pairs of symbols A_j occurring in G_i and $\nu_{i,j}$ the number of symbols R_j occurring in G_i. We consider G_1 before G_2, if G_1 is of less total degree in the coefficients of the two quadratics (2). But, if G_1 and G_2 are of the same degree in the coefficients of the two quadratics, we consider G_1 before G_2, if

$$\mu_{1,i+k} = \mu_{2,i+k}, \quad \nu_{1,i+k} = \nu_{2,i+k} \qquad (k = 1, 2, \cdots, n-i)$$

and

$$\mu_{1,i} > \mu_{2,i}, \ \nu_{1,i} \not< \nu_{2,i}, \ \text{or} \ \nu_{1,i} > \nu_{2,i}, \ \mu_{1,i} \not< \mu_{2,i}.$$

In other words, G_1 is considered before G_2, if the currency of G_1 is greater than the currency of G_2.

That this is a legitimate procedure follows from the fact, that each G_i may be treated as a concomitant K and may therefore be expressed in terms of the symbols $i, j, \cdots, k$, in which the symbols A_t, R_m convolved in G_i will also be convolved. For a similar reason, if in G_i a new convolution of symbols A or R is made, the complementary symbols may also be convolved together. Finally, if A_n or R_n appears in G_i, A_n^2 or R_n^2 is a factor of G_i and G_i is said to be reducible.

As a direct consequence of the order in which the concomitants G_i are considered, we may use formulae (14), (15) and lemma III in determining the G_i, since in lemma III the neglected terms are all of greater currency.

Bracket Factors with equivalent symbols. If Θ denote a convolution of symbols $i, j, \cdots, k$ and Φ a similar convolution, while no symbol i occurs

in Θ equivalent to a symbol j occurring in Φ and no symbol in Θ is a symbol successive to one in Φ, we may write

$$(\Theta\Phi \mid p) = (\Theta \mid \dot{X})(\Phi \mid \dot{Y})$$

without disturbing the invariant factors D_i, which may have been introduced by (14) and (15). Hence in considering factors with equivalent symbols we need only consider factors of the one type

$$(i, i+1, \cdots, i+k, t, s, \cdots, m \mid X) \qquad (i \leqq t, s, m \leqq i+k).$$

If the symbols t, s, m are not successive integers, we may break the factor up again into a series of products of factors without disturbing the invariant factors D_i. Therefore all factors involving equivalent symbols may be reduced to the type

$$T = (i, i+1, \cdots, i+k, t, t+1, \cdots, t+s \mid X)$$
$$(i \leqq t;\ i+k \geqq t+s),$$

where of course the variables p have been decomposed. Writing for i its value $(A_i R_{n+1-i})$ we see that T involves a convolution of

$$k+1+s+1 = k+s+2$$

equivalent symbols a. Since A_{i+k} is the symbol A of greatest currency appearing in T and since none of the symbols R have been disturbed, if $k+s+2 > i+k$, the currency of the resulting G_i has been increased. Hence, according to our scheme of order, G_i is reducible, if G_i contain the factor T, in which $s+2 > i$. We now consider the case in which $s+2 \leqq i$. If I denote the invariant factors, which appear in T by lemma III,

$$T \equiv I(\dot{A}_{i-1} R_{n+1-i})(\bar{B}_{t-1} S_{n+1-t})(\dot{A}_{k+1} \bar{B}_{s+1} \mid X_{k+s+2}),$$
$$\equiv I(\bar{B}_{s+1} \dot{A}_{i-2-s} R_{n+1-i})(\bar{B}_{t-1} S_{n+1-t})(\dot{A}_{k+s+2} \mid X_{k+s+2})$$

mod increased currency by (13), since $i-1 \geqq s+1$. But $n+1-i \geqq n+1-t$ and by transferring B [illegible] from the first to the second factor we displace some of S [illegible] and so increase the currency of R [illegible]. This [illegible]

are interchanged, the invariant factors D_i will not be disturbed. We have now proved the theorem:

THEOREM I. *Every concomitant K of the two quadratics (2) is reducible or else, if multiplied by an invariant factor, composed of powers of $(A_i R_{n-i})^2$ $(i=1, 2, \cdots, n-1)$, is a sum of terms, where each term is a product of the symbolic factors i_x, $(ij \mid p_2)$, $(ijk \mid p_3), \cdots, (ijk \cdots m \mid p_{n-1})$, $(12 \cdots n)$ where $i, j, k, m = 1, 2, \cdots, n$, and each symbol i occurs an even number of times in each product, while no equivalent symbols occur in the same factor. If the concomitant K is reducible, $K = IJ$, where I is an invariant factor composed of powers of $(A_n)^2$ and $(R_n)^2$ and J is not reducible.*

Since the variables are now convolved back again into the form p_j, we may replace the variables p by the variables π, thus removing the powers of D, with which we originally multiplied the concomitants. In fact, since the co-ordinates are homogeneous, it is immaterial which set of variables is used, as long as the variables are properly convolved. Henceforward we shall drop the variables and write (ij) for $(ij\pi_2)$, (ijk) for $(ijk\pi_3)$ etc.

4. *Identities.* If I_m, J_m, K_m etc. denote convolutions of m symbols $i, j, k, \cdots$ the following identities exist:

if $i + j + k = i + r + m$,

$$(16)\quad \begin{cases} (I_i \dot{J}_j K_k)(I_i \dot{J}_r M_m) \equiv (I_i J_{j+r} \dot{K}_{k-r})(I_i \dot{K}_r M_m) & (k > r), \\ \qquad\qquad \equiv (I_i J_{j+r})(I_i K_k M_m) & (k = r), \\ \qquad\qquad \equiv 0 & (k < r). \end{cases}$$

For, if the variables are taken as p, in transferring J_r to the first factor we cannot displace any of the variables, since otherwise the second factor would be zero. But, if $i + j + k = i + r + m - 1$,

$$(17)\quad \begin{cases} (I_i \dot{J}_j K_k)(I_i \dot{J}_r M_m) \equiv (I_i J_{j+r} \dot{K}_{k-r})(I_i \dot{K}_r M_m) + (I_i J_{j+r} \dot{K}_{k-r+1})(I_i \dot{K}_{r-1} M_m) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (k > r-1), \\ \qquad\qquad \equiv (I_i J_{j+r})(I_i K_{r-1} M_m) \qquad\qquad (k = r-1), \\ \qquad\qquad \equiv 0 \qquad\qquad (k < r-1). \end{cases}$$

The two terms in this identity occur, because one of the variables in the first factor can be displaced. Other identities exist, when $i+j+k<i+r+m-1$,

but they disturb the proper convolution of the variables. These identities (16) and (17) are true whether the variables p or π are used.*

5. *Principle of Duality.* Let K be a concomitant $\Pi(ij \cdots k)$. Then, if $i', j', \cdots, k'$ are the symbols complementary to $i, j, \cdots, k$, that is, if $(ij \cdots k)(i'j' \cdots k')$ involves each of the symbols $1, 2, \cdots, n$ once and only once, $(ij \cdots k)$ and $(i'j' \cdots k')$ are said to be dual factors. With this notation, $\Pi(ij \cdots k)\Pi(i'j' \cdots k')$ involves all the symbols $1, 2, \cdots, n$ exactly m times if there are m factors in K. If m is even, all the symbols occur an even number of times and therefore $\Pi(i'j' \cdots k')$ is also a concomitant. But, if m is odd $(12 \cdots n)\Pi(i'j' \cdots k')$ is a concomitant. Hence from a concomitant K we can form a dual concomitant by writing the dual factor of each factor in K and multiplying by the factor $(12 \cdots n)$, if the number of factors in K is odd. If we consider identity (16) and write the dual of each factor in it, we get,

$$(TM_m\dot{J}_r)(T\dot{J}_jK_k) \equiv (T\dot{K}_rM_m)(TJ_{j+r}\dot{K}_{k-r}),$$

where $(TM_mJ_{j+r}K_kI_i) = (12 \cdots n)$, and this new identity permutes exactly the same symbols as did the original one. Similarly there is an identity dual to (17) acting on the same symbols as (17). Hence we have the important lemma:

LEMMA IV. *Corresponding to any identity there is a dual identity operating on the same symbols.*

Since, if i and j are successive integers and ij is convolved an even number of times in a concomitant, an identity separating i and j cannot be used without disturbing the invariant factors, we must now prove a further lemma.

LEMMA V. *If K and K' are dual concomitants, the symbols i and j are convolved an even or an odd number of times in K' according as they are convolved an even or an odd number of times in K.*

Let the symbols i and j occur alone in m and r factors of K respectively, the symbols i, j occur together in t factors of K, and neither of the symbols i or j occur in q factors of K. Then the total number of factors in K is $m + r + t + q = s$. Since $t + m$ and $t + r$ are both even, $s + t \equiv q + m + t + r + t \equiv q \bmod 2$, or $q \equiv t - s \bmod 2$. *Hence*

$$t \equiv q \bmod 2, \text{ if } s \text{ is even,}$$
$$\equiv q + 1 \bmod 2, \text{ if } s \text{ is odd.}$$

* Turnbull, *loc. cit.*, pp. 92 sq.

But if s is even, q is the number of factors in K' containing both i and j, while, if s is odd, $q+1$ is the number of factors in K' containing both i and j, since, if s is odd, in determining K' from K we introduce the factor $(12\cdots n)$. Accordingly this proves our lemma.

6. *Incompleteness of the Prepared System.* We should like to determine a prepared system, in terms of which every concomitant can be expressed, without being multiplied by an invariant factor. To do this, we must see, if ever in convolving the variables to give the factors $(ij\cdots k)$ we have broken up a convolution of successive symbols. This is the case for $n>3$, and accordingly we must introduce other more complicated factor types. With the introduction of these new factor types, once we have found a complete system for the concomitants multiplied by invariant factors, we may remove the invariant factors from each concomitant and so obtain the actual complete system, provided that no identity is used, which disturbs a convolution of two successive symbols, if that convolution occurs an even number of times in a concomitant. These new factor types are of the nature

$$(\ddot{i}\dot{j}\cdots t)(\dot{k}\cdots\bar{s})(\bar{m}\cdots)(\bar{n}\cdots)\cdots(\bar{r}\cdots),$$

or in other words combinations of simple bracket factors, in which successive symbols are implicitly convolved. The dual of such a factor is obtained by taking the duals of the component bracket factors and permuting the same symbols. Accordingly, if one of these new factor types is counted as equivalent to q simple factors, where q is the number of component factors in the new factor type, the dual of a concomitant K containing such a factor can be written down as in § 4. Further the introduction of these new factor types does not vitiate the proof of lemma V, since any symbols implicitly convolved in K will also be implicitly convolved in K'. Hence from a list of irreducible concomitants we can immediately write down a list of the dual irreducible concomitants.

Since the concomitants f and g enter symmetrically into the discussion, the actual labour of determining the irreducible concomitants can be shortened still further. Two factors, concomitants or identities are said to be *similar,* if the first can be derived from the second by writing $n+1-i$ for every symbol i that occurs. From the symmetry mentioned above, it follows immediately that, if a concomitant is reducible, so is the concomitant similar to it.

7. *Applications of the General Theory to the Cases $n=2$ and 3. Complete System for Two Quadratics in 2 Variables.* In this case $n=2$ and we

require two variables, only one of which occurs explicitly. It is obvious that the prepared system consists of the three simple factors:

$$1_x = a_x, \quad 2_x = r_x, \quad (12) = (ar).$$

The prepared system then consists of six forms:

three invariants $(A_2)^2$, $(R_2)^2$, $(ar)^2$;
three covariants 1_x^2, 2_x^2, $(12)1_x 2_x$.

In addition there is the actual concomitant of the field $D = (uv)$.*

Complete System for Two Quadratics in 3 *Variables.* In this case $n = 3$ and we require two co-ordinates x and $u = (xy)$. The prepared system consists of the seven simple factors:

$$1_x = a_x, \quad 3_x = r_x, \quad 2_x = (ARx) = (aR)b_x - (bR)a_x,$$
$$(12) = (aR)(Au), \quad (32) = (rA)(Ru), \quad (13) = (aru),$$
$$(123) = (aR)(Ar),$$

where A and R are written for A_2 and R_2 respectively. The complete system consists of 20 forms:

four invariants $(A_3)^2$, $(Ar)^2$, $(aR)^2$, $(R_3)^2$;
four covariants 1_x^2, 2_x^2, 3_x^2, $(123)1_x 2_x 3_x$;
four contravariants $(12)^2$, $(23)^2$, $(13)^2$, $(23)(31)(12)$;
eight mixed forms $(123)(23)1_x$, $(123)(12)3_x$, $(ij)i_x j_x$, $(123)(ij)(ik)i_x$,

where $i, j, k = 1, 2, 3$ and $i \neq j \neq k$. In addition there is the actual concomitant of the field $D = (uvw)$. The actual concomitants are obtained from the above list by dropping the invariant factors, e. g. $(23)^2 = (rA)^2(Ru)^2$ gives the actual concomitant $(Ru)^2$.†

Determination of the Complete System. No new types of bracket factors are necessary, since the only chance of 12 or 23 being separated would be in the formation of the factor (123). But, as in (123) both 12, and 23 are convolved, no invariant factor would be disturbed. Accordingly the prepared system is that quoted above.

Covariants. If the factor (123) does not occur, the only possibilities are the squares of the factors i_x $(i = 1, 2, 3)$. If the factor (123) does occur, the only possibility is $(123)1_x 2_x 3_x$.

---

* Weitzenböck, *Invariantentheorie*, Chap. 2, pp. 42, 43.

† Weitzenböck, *loc. cit.*, Chap. 2, page 61.

Contravariants. By the principle of duality we obtain the contravariants immediately from the list of covariants.

Mixed Concomitants. If the factor (123) does not occur, by analogy with binary forms the only possibilities are the three forms $(ij)i_x j_x$ $(i, j = 1, 2, 3)$. From these we obtain the dual forms $(123)(kj)(ki)k_x$ $(i, j, k = 1, 2, 3)$. If the factor (123) occurs, we may have the forms

$$(123)(23)1_x, \quad (123)(21)3_x, \quad (123)(13)2_x,$$

of which the last can be expressed in terms of the other two. It is obvious that three u factors cannot occur, since then we would have the concomitant

$$(123)(12)(23)(31)1_x 2_x 3_x$$

and this has the concomitant factor $(13)1_x 3_x$. Hence we have obtained the complete list of irreducible concomitants.

Similar methods have been employed to find the complete system for the case when $n = 4$.* When $n = 5$ and 6, an actual prepared system has been found by these methods, but the work is too long to be reproduced here.

THE JOHNS HOPKINS UNIVERSITY.

---

* J. Williamson, "The Complete System of Two Quarternary Quadratics," *American Journal of Mathematics*, Vol. 51, No. 4 (October, 1929). In this paper Theorem 1 is proved solely for the case $n = 4$.

On Algebraic Inversive Invariants.

By Frank Morley and Boyd C. Patterson.

Introduction. The invariant theory of the inversive group has been investigated by Kasner * most thoroughly. His mode of approach was by the study of "the theory of algebraic curves upon a proper quadric surface, with respect to those properties which are unaltered by the group of collineations transforming the quadric into itself." When the quadric is a sphere this group of collineations affects the points on the sphere in the same way as the group of inversions upon the sphere; and by an inversion of the sphere into a plane the geometry of the above mentioned theory of algebraic curves reduces to the inversive geometry of the plane.

It is the purpose of this paper to present new methods for calculating the algebraic invariants of curves under the inversive group. In particular, the invariants and covariants of the biquadratic are obtained. The method employed rests on the inversive significance of the equation of a curve given in circular coördinates. A bilinear equation represents a circle or a straight line (inversively equivalent to a circle) or two points, images in the plane considered. The biquadratic equation includes all circles which, with a circle, determine four points. Thus the curves otherwise called the conics, circular cubics, bicircular quartics, Cartesians, and Cassinians are all, inversively, biquadratics.

It is worthy of note that the methods presented may easily be generalized; and the results of the application of this method to the bilinear and bicubic forms are given.

In Section 4 an application of the results obtained for the general biquadratic is made to the Neuberg curve of a triangle. Sometimes called the Neuberg cubic (projectively speaking), from our point of view it is a biquadratic through infinity.

1. *Concomitants of the biquadratic.* We write the equation of the biquadratic

$$(1.1)\qquad (a_{22}x^2+2a_{21}x+a_{20})\bar{x}^2+2(a_{12}x^2+2a_{11}x+a_{10})\bar{x} + (a_{02}x^2+2a_{01}x+a_{00})=0$$

in circular coördinates with complex coefficients, $x=X-iY$ representing

* Edward Kasner, "The Invariant Theory of the Inversion Group: Geometry upon a Quadric Surface," *Transactions of the American Mathematical Society*, Vol. 1 (1900), p. 430.

a point of the complex plane and $\bar{x} = X - iY$ its reflexion in the axis of reals, i. e. its conjugate. An equation $f(x, \bar{x}) = 0$ represents a real curve only when it is self-conjugate. Hence we impose the condition that $\bar{a}_{ij} = a_{ji}$ from which it follows that a_{ii} is real. The coördinates may be modified by a real factor; thus we speak of the form itself as the curve.

In the sequel we shall frequently write the left member of (1. 1) in the following matrix form:

$$(1.1') \qquad \begin{array}{c|ccc} C_{12} & 1 & 2x & x^2 \\ \hline 1 & a_{00} & a_{01} & a_{02} \\ 2\bar{x} & a_{10} & a_{11} & a_{12} \\ \bar{x}^2 & a_{20} & a_{21} & a_{22}. \end{array}$$

By the notation $f \equiv C_{mn}$ we shall indicate that f is a form of degree m in the coefficients and of degree n in each of the variables x and $\bar{x}$. Equation (1. 1) gives points on the biquadratic; a more general form of the equation is obtained by replacing $\bar{x}$ by $\bar{y}$. It is known from the general theory that an equation $f(x, \bar{y}) = 0$ does not represent a curve, since it is not self-conjugate, but rather associates with every point x of the plane certain points y. These points y are called image points of x with respect to the curve. In the case at hand we notice that for a given x we have a quadratic in $\bar{y}$ —hence two images y.

It can be shown with little difficulty that a $C_{12} = 0$ is transformed into a curve of the same type under transformations of the kind $y = \alpha x + \beta$ (a proportion or homology) and $z = 1/\bar{y}$ (an inversion in the base circle). Since the group of inversions consists of homologies and inversions in the unit circle, it follows that the type of the curve is left unaltered by the transformations of the inversive group

$$(1.2) \qquad y = (ax + b)/(cx + d), \qquad z = 1/\bar{y}.$$

The form C_{12} of the curve has therefore an inversive invariant significance. Furthermore the determinant of the coefficients is invariant under inversions. That is, if $C_{12} = 0$ is transformed into $C'_{12} = 0$ by a transformation (1. 2) the expression

$$I_3 = \begin{vmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{vmatrix}$$

is, to within a constant factor, equal to the similar expression I'_3 of C'_{12}. In fact $I'_3 = (M\bar{M})^3 I_3$, $M = ad - bc$ being the modulus of the transformation. We note that the above matrix is an Hermitian matrix since $\bar{a}_{ij} = a_{ji}$.

Knowing I_3 to be an invariant of C_{12} we may calculate other invariants as follows. Writing down the invariant I_3 of the pencil of biquadratics

$$C_{12} + kP \qquad k, \text{ real}$$

and expanding in powers of k, we have invariants of the system appearing as coefficients of the various powers of k. If now P should be a degenerate biquadratic, say a point z taken four times,

$$P \equiv (x - z)^2 (\bar{x} - \bar{z})^2 = 0,$$

we find the coefficients of k to be invariants of the system $C_{12} + kP$ if z is considered a constant or covariants of C_{12} if z is taken to be a variable. The invariant of the pencil $C_{12} + kP$ reduces to

$$kC_{22} + I_3.$$

The coefficient of k is a covariant of the original form of degree two in the coefficients a_{ij}—it is the first evectant* of C_{12}. Changing the variable from z to x, the matrix form of C_{22} is

	1	x	x^2
1	A_{22}	A_{21}	A_{20}
$\bar{x}$	A_{12}	A_{11}	A_{10}
$\bar{x}^2$	A_{02}	A_{01}	A_{00}

where A_{ij} is the minor of a_{ij} in I_3. We notice the binomial coefficients to be lacking in the C_{22} although they were present in C_{12}. If now we form the first evectant of the pencil $C_{22} + kP$ we find it to be precisely $I_3 C_{12}$. This reciprocal relation between the two forms C_{12} and C_{22} leads us to speak of C_{22} as the reciprocal of C_{12}.

The invariants of the system

$$C_{12} + kC_{22}$$

will be invariants of the form C_{12}. We obtain these invariants as coefficients of powers of k in the expansion of the following determinant which is the invariant I_3 of the pencil $C_{12} + kC_{22}$:

$$\begin{vmatrix} a_{00} + kA_{22} & 2a_{01} + kA_{21} & a_{02} + kA_{20} \\ 2a_{10} + kA_{12} & 4a_{11} + kA_{11} & 2a_{12} + kA_{10} \\ a_{20} + kA_{02} & 2a_{21} + kA_{01} & a_{22} + kA_{00} \end{vmatrix} = I_3{}^2 k^3 + 2I_3 I_2 k^2 + I_4 k + 4I_3.$$

* Cf. Salmon, *Modern Higher Algebra*, 2nd edition (1866), p. 105.

The values of these new invariants are:

$$I_2 = 2a^2_{11} + a_{00}a_{22} + a_{20}a_{02} - 2a_{10}a_{12} - 2a_{01}a_{21}.$$
$$I_4 = A^2_{11} + 8A_{00}A_{22} + 8A_{02}A_{20} - 4A_{10}A_{12} - 4A_{01}A_{21}.\text{*}$$

The covariants are obtained as coefficients of the powers of k in the development of the bordered determinant:

$$\begin{vmatrix} 0 & x^2 & -2x & 1 \\ \bar{x}^2 & a_{00} + kA_{22} & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot \\ -2\bar{x} & \cdot\;\cdot\;\cdot & 4a_{11} + kA_{11} & \cdot\;\cdot\;\cdot \\ 1 & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & a_{22} + kA_{00} \end{vmatrix}$$
$$= -I_3C_{12}k^2 - (I_2C_{12} - 2C_{32})k - 4C_{22}.$$

The covariants C_{12} and C_{22} are the original form and its reciprocal. The third, a C_{32}, is as follows:

C_{32}	1	$2x$	x^2
1	α_{00}	α_{01}	α_{02}
$2\bar{x}$	α_{10}	α_{11}	α_{12}
$\bar{x}^2$	α_{20}	α_{21}	α_{22}

where the coefficients α_{ij} have the following values:

$$\alpha_{22} = a_{22}(a_{02}a_{20} - a^2_{11}) + 2a_{11}a_{21}a_{12} - a^2_{12}a_{20} - a^2_{21}a_{02}.$$
$$\alpha_{12} = a_{01}(a_{12}a_{21} - a_{11}a_{22}) + a_{02}(a_{10}a_{22} - a_{11}a_{21}) - a_{12}(a_{10}a_{12} - a^2_{11}).$$
$$\alpha_{02} = a_{00}(a_{02}a_{22} - a^2_{12}) - a^2_{01}a_{22} + 2a_{01}a_{11}a_{12} - a_{02}a^2_{11}.$$
$$2\alpha_{11} = a_{01}(a_{12}a_{20} - a_{11}a_{21}) + a_{02}(a_{10}a_{21} - a_{11}a_{20}) - 2a_{11}(a_{10}a_{12} - a^2_{11})$$
$$- a_{00}(a_{11}a_{22} - a_{12}a_{21}) + a_{01}(a_{10}a_{22} - a_{21}a_{11}).$$
$$\alpha_{01} = a_{00}(a_{02}a_{21} - a_{12}a_{11}) - a_{01}(a_{21}a_{01} - a^2_{11}) + a_{10}(a_{01}a_{12} - a_{11}a_{02}).$$
$$\alpha_{00} = a_{00}(a_{02}a_{20} - a^2_{11}) - a^2_{01}a_{20} + 2a_{01}a_{10}a_{11} - a^2_{10}a_{02}.$$
$$\bar{\alpha}_{ij} = \alpha_{ji}.$$

It is known that a biquadratic curve is its own inverse with respect to any one of its four Jacobian or "director" circles. Inversion with respect to an intersection of any two of these circles (which are orthogonal) sends the curve into a biquadratic symmetrically placed with respect to two perpendicular lines. Taking these lines as coördinate axes, the equation of the biquadratic assumes the normal form:

$$C_{12} \equiv \kappa(x^2\bar{x}^2 + 1) + \lambda(x^2 + \bar{x}^2) + 4\mu x\bar{x}.$$

The normal forms of the other covariants and invariants are:

$$C_{22} \equiv \kappa\mu(x^2\bar{x}^2 + 1) - \lambda\mu(x^2 + \bar{x}^2) + (\kappa^2 - \lambda^2)x\bar{x},$$
$$C_{32} \equiv \kappa(\lambda^2 - \mu^2)(x^2\bar{x}^2 + 1) + \lambda(\kappa^2 - \mu^2)(x^2 + \bar{x}^2) - 2\mu(\kappa^2 + \lambda^2 - 2\mu^2)x\bar{x},$$
$$I_2 = \kappa^2 + \lambda^2 + 2\mu^2, \quad I_3 = \mu(\kappa^2 - \lambda^2), \quad I_4 = (\kappa^2 - \lambda^2)^2 + 8\mu^2(\kappa^2 + \lambda^2).$$

---

* Kasner, *loc. cit.*, p. 479, gives the invariants I, J, K. In terms of the above, $I = -I_2$, $J = 2I_3$, $4K = I_2^2 - I_4$.

The above normal forms rule out the conics as special biquadratics. They suffice, however, to calculate important relations between the invariants.

Writing C_{22} with binomial coefficients as in (1. 1′) we have the following relations between the invariants J of this form and the invariants I of C_{12}:

$$8J_2 = I_4, \qquad 4J_3 = I_3^2, \qquad 2J_4 = I_3^2 I_2.$$

Repeating the process on the reciprocal of C_{22} and indicating by K the invariants of this form, we have

(1. 3) $\quad 16K_2 = 2J_4 = I_3^2 I_2, \quad 64K_3 = 16J_3^2 = I_3^4, \quad 256K_4 = 128J_3^2 J_2 = I_3^4 I_4.$

Foci. The image equation of the general *bi-n-ic* associates with any point x of the plane n image points y. By definition a focus of the curve is a point x which has two coincident images. Writing the biquadratic C_{12} in the form

$$(a_{22}x^2 + 2a_{21}x + a_{20})\bar{y}^2 + 2(a_{12}x^2 + 2a_{11}x + a_{10})\bar{y} + (a_{02}x^2 + 2a_{01}x + a_{00}) = 0$$

and requiring the discriminant of this quadratic in $\bar{y}$ to vanish makes evident the fact that $C_{12} = 0$ has four foci. The foci of the normal form are given by the quartic

$$Q \equiv \kappa\lambda(x^4 + 1) + (\kappa^2 + \lambda^2 - 4\mu^2)x^2 = 0.$$

Writing the biquadratic as a quadratic in x and requiring its discriminant to vanish gives four points y as foci which are identical to the four roots of Q. The projective invariants and covariants of Q are:

$$g_2 = \kappa^2\lambda^2 + 3A^2, \qquad g_3 = A(\kappa^2\lambda^2 - A^2), \qquad \text{where } 6A = \kappa^2 + \lambda^2 - 4\mu^2;$$

$$\Delta_{12} \equiv g_2^3 - 27g_3^2$$
$$= \{(1/4)\kappa\lambda(\kappa + \lambda + 2\mu)(\kappa + \lambda - 2\mu)(\kappa - \lambda + 2\mu)(-\kappa + \lambda + 2\mu)\}^2,$$

$$H \equiv \kappa\lambda A(x^4 + 1) + (\kappa^2\lambda^2 - 3A^2)x^2, \qquad S \equiv \kappa\lambda(\kappa^2\lambda^2 - 9A^2)(x^5 - x),$$

where the subscripts of Δ indicate its degree in the coefficients of C_{12}; H is the Hessian and S the sextic covariant. Foci invert into foci, hence we expect the above invariants of Q to be expressible in terms of the inversive invariants of C_{12}. In fact we have

$$12g_2 = 4I_2^2 - 3I_4, \qquad 216g_3 = 8I_2^3 + 54I_3^2 - 9I_2I_4,$$

$$\Delta_{12} = (1/64)(I_2^2I_4^2 - I_4^3 - 108I_3^4 - 32I_2^3I_3^2 + 36I_2I_3^2I_4).$$

The foci of C_{22} are given by

$$Q' \equiv -\kappa\lambda\mu^2(x^4 + 1) - \{(1/4)(\kappa^2 - \lambda^2)^2 - \mu^2(\kappa^2 + \lambda^2)\}x^2 = 0.$$

Between Q, Q', and H there exists the relation

$$Q' = -(I_2/6)Q + H.$$

Q' then is one of the pencil of quartics $uQ + vH$ and the invariants of Q' may be expressed in terms of the invariants of Q.* In fact

* Salmon, *loc. cit.*, p. 177.

$$(1.4)\qquad \begin{aligned} 36g'_2 &= I_2{}^2 g_2 - 18 I_2 g_3 + 3 g_2{}^2, \\ 216 g'_3 &= - I_2{}^3 g_3 + I_2{}^2 g_2{}^2 - 9 I_2 g_2 g_3 + 54 g_3{}^2 - g_2{}^3, \\ 256\Delta'_{24} &= I_3{}^4 \Delta_{12}, \qquad -16 S' = I_3{}^2 S. \end{aligned}$$

In addition we have the identity

$$4H^3 - g_2 H Q^2 + g_3 Q^3 + S^2 \equiv 0,$$

and the invariants g_2 and g_3 expressed in terms of g'_2 and g'_3:

$$\begin{aligned} 3^2 I_3{}^4 g_2 &\doteq I_4{}^2 g'_2 - 2^4 \cdot 3^2 I_4 g'_3 + 2^6 \cdot 3 g'_2{}^2 \\ 3^3 I_3{}^6 g_3 &= - I_4{}^3 g'_3 + 8 I_4{}^2 g'_2{}^2 - 8^2 \cdot 3^2 I_4 g'_2 g'_3 + 8^3 \cdot 2 \cdot 3^3 g'_3{}^2 - 8^3 g'_2{}^3. \end{aligned}$$

Two Biquadratics. We close this section with a word concerning the simultaneous invariants of two biquadratics C_{12} and C'_{12} with coefficients a_{ij} and b_{ij} respectively. Forming the pencil

$$C_{12} + kC'_{12}$$

and calculating the invariants corresponding to I_2, I_3, and I_4 of C_{12}, we obtain other invariants involving the coefficients of both forms. For future reference we give here the bilinear or apolar invariant of the two forms:

$$\begin{aligned} I_{11} = 4a_{11}b_{11} + a_{00}b_{22} + a_{22}b_{00} + a_{20}b_{02} + a_{02}b_{20} \\ - 2(a_{10}b_{12} + a_{12}b_{10} + a_{01}b_{21} + a_{21}b_{01}). \end{aligned}$$

2. *Invariants of circles.* Applying the method of the preceding section to the bilinear forms

$$\begin{aligned} C_{11} &\equiv \alpha_{11} x\bar{x} + \alpha_{01} x + \alpha_{10}\bar{x} + \alpha_{00} \\ C'_{11} &\equiv \beta_{11} x\bar{x} + \beta_{01} x + \beta_{10}\bar{x} + \beta_{00} \end{aligned}$$

gives us the invariant

$$I = \alpha_{00}\alpha_{11} - \alpha_{10}\alpha_{01}$$

of the single form, and

$$J = \alpha_{11}\beta_{00} - \alpha_{01}\beta_{10} - \alpha_{10}\beta_{01} + \alpha_{00}\beta_{11},$$

the simultaneous invariant of the two forms.

Geometrically interpreted, $C_{11} = 0$ is a circle which is a point when $I = 0$. The vanishing of J is the condition that the two circles be orthogonal.

If the four circles

$$\begin{aligned} f_1 &\equiv \alpha_{11} x\bar{x} + \alpha_{01} x + \alpha_{10}\bar{x} + \alpha_{00}, \qquad & f_2 &\equiv \beta_{11} x\bar{x} + \beta_{01} x + \beta_{10}\bar{x} + \beta_{00}, \\ f_3 &\equiv \gamma_{11} x\bar{x} + \gamma_{01} x + \gamma_{10}\bar{x} + \gamma_{00}, \qquad & f_4 &\equiv \delta_{11} x\bar{x} + \delta_{01} x + \delta_{10}\bar{x} + \delta_{00}, \end{aligned}$$

have a common orthogonal circle

$$f \equiv a_{11} x\bar{x} + a_{01} x + a_{10}\bar{x} + a_{00}$$

then

$$J(f_k f) = 0 \qquad (k = 1, 2, 3, 4).$$

Eliminating a_{ij} from these four equations gives us the invariant condition that f_k have a common orthogonal circle:

$$K \equiv \begin{vmatrix} \alpha_{11} & \alpha_{01} & \alpha_{10} & \alpha_{00} \\ \beta_{11} & \beta_{01} & \beta_{10} & \beta_{00} \\ \gamma_{11} & \gamma_{01} & \gamma_{10} & \gamma_{00} \\ \delta_{11} & \delta_{01} & \delta_{10} & \delta_{00} \end{vmatrix} = 0.$$

Two circles $C_{11}=0$ and $C'_{11}=0$ have one absolute invariant,

$$J^2/II' = 4\cos^2\theta \text{ or } 4\cosh^2\delta,$$

where θ is the angle when the circles intersect; δ is the hyperbolic distance (measured from the common inverse points) when they do not. Hence two circles are tangent when

$$\begin{vmatrix} J & 2I \\ 2I' & J \end{vmatrix} = 0.$$

3. *Interpretations of the invariants of* C_{12}. Before entering upon the subject of this section we define several circles* intimately connected with the biquadratic C_{12}. The equations for the most part will be given in the normal form.

(α) The conjugate point x of a point z with respect to C_{12} is the harmonic conjugate of z with respect to the image points of z. It is given by the equation

$$(\kappa\bar{z}^2+\lambda)xz + 2\mu\bar{z}(x+z) + (\kappa+\lambda\bar{z}^2) = 0.$$

(β) The bipolar circle C' of a point z with respect to C_{12} is the locus of double points of the one-to-one correspondence

$$(\kappa\bar{y}\bar{z}+\lambda)xz + \mu(\bar{y}+\bar{z})(x+z) + (\kappa+\lambda\bar{y}\bar{z})$$

determined by polarizing the image equation of C_{12} with respect to x and $\bar{y}$ simultaneously. Its equation is

$$C' \equiv (\kappa z\bar{z}+\mu)x\bar{x} + (\lambda z+\mu\bar{z})x + (\lambda\bar{z}+\mu z)\bar{x} + (\kappa+\mu z\bar{z}) = 0.$$

We note that the conjugate point of z is the inverse of z with respect to C'.

(γ) The bipolar circle C_1 of a circle C with respect to C_{12} is the locus of points whose bipolar circles are orthogonal to C. If the circle C has the equation

$$C \equiv \alpha_{11}x\bar{x} + \alpha_{01}x + \alpha_{10}\bar{x} + \alpha_{00} = 0$$

the equation of its bipolar circle is

$$C_1 \equiv (\alpha_{11}\mu + \alpha_{00}\kappa)x\bar{x} - (\alpha_{01}\mu + \alpha_{10}\lambda)x \\ - (\alpha_{10}\mu + \alpha_{01}\lambda)\bar{x} + (\alpha_{11}\kappa + \alpha_{00}\mu) = 0.$$

* See Kasner, *loc. cit.*, pp. 482, 485 for definitions in quaternary form. Properly for circle we should write bilinear curve.

(δ) The anti-bipolar * circle C_2 of a circle C with respect to C_{12} is the circle orthogonal to the bipolar circles of all points on C. The bipolar circle of the anti-bipolar circle is the original circle. Hence the equation of the anti-bipolar circle of C is

$$C_2 \equiv (\lambda^2 - \mu^2)(\alpha_{00}\kappa - \alpha_{11}\mu)x\bar{x} + (\kappa^2 - \mu^2)(\alpha_{01}\mu - \alpha_{10}\lambda)x + (\kappa^2 - \mu^2)(\alpha_{10}\mu - \alpha_{01}\lambda)\bar{x} + (\lambda^2 - \mu^2)(\alpha_{11}\kappa - \alpha_{00}\mu) = 0.$$

(ϵ) When C reduces to a point z we have the anti-bipolar circle C'' of z with respect to C_{12}. Its equation is

$$C'' \equiv (\lambda^2 - \mu^2)(\kappa z\bar{z} - \mu)x\bar{x} + (\kappa^2 - \mu^2)(\lambda z - \mu\bar{z})x + (\kappa^2 - \mu^2)(\lambda\bar{z} - \mu z)\bar{x} + (\lambda^2 - \mu^2)(\kappa - \mu z\bar{z}) = 0.$$

A geometrical interpretation of the covariants is obtained by the application of the results of Section 2 to the bipolar circles P_{12} and P_{22} of a point with respect to C_{12} and C_{22} respectively. Since $I(P_{22}) = (I_3/4)C_{12}$, $C_{12} = 0$ is the locus of points whose bipolar circles with respect to C_{22} are degenerate. Also $C_{12} = 0$ is the original biquadratic and the bipolar circle with respect to C_{12} of each point on it passes through that point. $I(P_{12}) = C_{22}$, hence $C_{22} = 0$ is the locus of points whose bipolar circles with respect to C_{12} are degenerate. Finally, $C_{32} = 0$ is the locus of a point whose anti-bipolar circle with respect to C_{12} passes through it.

Calculating the bilinear or apolar † invariant of the biquadratics C_{12}, C_{22}, and C_{32} gives the following facts:

$$I_{11}(C_{12}C_{12}) = 2I_2, \qquad I_{11}(C_{22}C_{22}) = I_4/4,$$
$$I_{11}(C_{12}C_{22}) = 3I_3, \qquad I_{11}(C_{12}C_{32}) = I_2{}^2 - I_4.$$

Hence $I_2 = 0$ is the condition that the biquadratic is self-apolar. $I_3 = 0$ is the condition that the biquadratic is apolar to its reciprocal curve C_{22}. And when $I_4 = 0$ the covariant curve C_{22} is self-apolar. We give here another interpretation of the vanishing of I_3.

Writing $C_{12} = 0$ as an image equation and polarizing with respect to x and $\bar{y}$ defines a one-to-one correspondence between two points x and y:

$$[a_{22}xx_1 + a_{21}(x + x_1) + a_{20}]\bar{y}\bar{y}_1 + [a_{12}xx_1 + a_{11}(x + x_1) + a_{10}](\bar{y} + \bar{y}_1) + [a_{02}xx_1 + a_{01}(x + x_1) + a_{00}] = 0.$$

Now the condition that there be a neutral pair, i. e., a pair of points x and x_1 for which y is arbitrary, is that the three equations

$$a_{i2}xx_1 + a_{i1}(x + x_1) + a_{i0} = 0 \qquad (i = 0, 1, 2)$$

have a solution, that is $I_3 = 0$.

---

* J. L. Coolidge, *Treatise on the Circle and the Sphere*, Oxford (1916), p. 211, calls this the auto-polar circle.

† Kasner, *loc. cit.*, p. 471.

$I_3 = 0$ is the condition that C_{12} has a neutral pair. Neutral pairs in the above sense are characteristic of Cassinians and their inverses.

We have defined a focus as a point whose images with respect to C_{12} are coincident. A biquadratic will have a double point if two foci coincide and the double point then absorbs the foci. Three coincident foci give the biquadratic a cusp. From the invariants of Q (Section 1) we have:

$C_{12} = 0$ is a nodal biquadratic when $\Delta_{12} = 0$. A nodal biquadratic is a central conic or the inverse of a central conic. The conic is an ellipse if the node is isolated, an hyperbola if the node is an ordinary double point.

$C_{12} = 0$ is a cuspidal biquadratic when $4H^3 + S^2 \equiv 0$, for $g_2 = 0$ and $g_3 = 0$. A cuspidal biquadratic is a parabola or the inverse of a parabola.

$C_{12} = 0$ has two nodes and degenerates into a pair of circles when $S \equiv 0$. For then Q has two pairs of double roots and H is, to a constant multiple, the same as Q; hence the Jacobian S of H and Q vanishes identically.

The relations (1. 3) and (1. 4) give us corresponding results concerning the reciprocal curve $C_{22} = 0$ and make the reciprocity of C_{12} and C_{22} even more striking.

4. *The Neuberg curve of a triangle.** Given three points a_1, a_2, a_3, their Neuberg curve is defined to be the locus of a point a_4 such that, if λ_{ij} is the square of the distance between a_i and a_j,

$$\Delta' \equiv \begin{vmatrix} \lambda_{23}\lambda_{14} & \lambda_{23} + \lambda_{14} & 1 \\ \lambda_{31}\lambda_{24} & \lambda_{31} + \lambda_{24} & 1 \\ \lambda_{12}\lambda_{34} & \lambda_{12} + \lambda_{34} & 1 \end{vmatrix} = 0.$$

The Neuberg curve has the property that the isogonal conjugate with respect to the triangle $a_1a_2a_3$ of each point on the curve also lies on the curve and their join is parallel to the Euler line of the triangle. This definition of the curve enables us to derive its equation in circular coördinates. For we can take the vertices of the triangle on the unit circle, say, t_1, t_2, t_3, and choose the axes so that the axis of imaginaries goes through the orthocenter. With this choice the circumcenter of the triangle is the origin and the orthocenter $s_1 = t_1 + t_2 + t_3$ is a pure imaginary, say $s_1 = 2\rho i$ where ρ is real and positive. The transformation of isogonal conjugates for this triangle is

If the join of a pair of isogonal conjugates x, y is to be parallel to the Euler line we must have

$$(x-y)/(\bar{x}-\bar{y})=-1.$$

The locus of isogonal conjugates x, y satisfying

$$x+y+s_3\bar{x}\bar{y}=2\rho i, \qquad \bar{x}+\bar{y}+xy/s_3=-2\rho i, \qquad x+\bar{x}=y+\bar{y}$$

is the Neuberg curve of the triangle $t_1t_2t_3$. Eliminating 1, y, $\bar{y}$, from these three equations, we find the locus to be

$$(4.1) \qquad x^2\bar{x}+x\bar{x}^2+x^2/s_3+s_3\bar{x}^2-2(1+\rho i/s_3)x-2(1-\rho is_3)\bar{x}=0.$$

We notice that the Neuberg curve is a biquadratic through infinity. (That this equation is the same as $\Delta'=0$ may be shown by direct substitution of t_i for a_i and of x for a_4).

A third characterization of the Neuberg curve is: The locus of points x whose image triangle with respect to the triangle $t_1t_2t_3$ is perspective to $t_1t_2t_3$. The equation of the curve in circular coördinates may also be derived from this definition.

It has been shown that the Neuberg curve is a symmetrical relation on five points,* and the property mentioned in the preceding paragraph tells what it is. Take any one of the five points, say α_4, and its three inverse points with respect to the three circles determined by $\alpha_2\alpha_3\alpha_5$, $\alpha_3\alpha_1\alpha_5$, and $\alpha_1\alpha_2\alpha_5$. Call these inverse points β_1, β_2, β_3, respectively. The distinctive property of the five points $\alpha_1,\cdots,\alpha_5$ is that the circles $\alpha_1\beta_1\alpha_5$, $\alpha_2\beta_2\alpha_5$, $\alpha_3\beta_3\alpha_5$, have in addition to α_5 a second common point, i. e. they are coaxal.

The Neuberg curve therefore is the locus of a point α_4 having the above relation with respect to four special points. The symmetry of $\Delta'=0$ in its inverted form shows that what is true inversively of one of the points is true of the others. Before we attempt to characterize these points inversively, i. e. by means of inversive invariants and covariants, we put them into a more convenient position by inverting with respect to one of them, say α_5. The four points now become the vertices of a triangle, which we call t_1, t_2, t_3, and the infinite point of the plane; and the equation of the curve is (4.1) which we shall indicate hereafter by C_{12}.

Applying the results of Section 1 we have the covariants

C_{12}	1	$2x$	x^2
1	0	$-(1+\rho i/s_3)$	$1/s_3$
$2\bar{x}$	$-(1-\rho is_3)$	0	$1/2$
$\bar{x}^2$	s_3	$1/2$	0

* Frank Morley, "Note on Neuberg's Cubic Curve," *American Mathematical Monthly*, Vol. 32 (1925), p. 407.

C_{22}	1	x	x^2
1	$-(1+2\rho\sigma+\rho^2)$	$(1/s_3-\rho i)$	$-(1+\rho i/s_3)/2$
$\bar{x}$	$(s_3+\rho i)$	-1	$-1/2s_3$
$\bar{x}^2$	$-(1-\rho i s_3)/2$	$-s_3/2$	$-1/4$

C_{32}	1	$2x$	x^2
1	$2\gamma(\rho^2-1)$	$\rho\gamma(\rho/s_3-i)$	0
$2\bar{x}$	$\rho\gamma(\rho s_3+i)$	$-\gamma/2$	$-\rho i\gamma/2$
$\bar{x}^2$	0	$\rho i\gamma/2$	$-\gamma/2$

where s_3 is a complex number of absolute value unity and has been written, when convenient, $\cos\phi + i\sin\phi = \gamma + i\sigma$; $\phi = \operatorname{amp} s_3$ and hence

$$2i\sigma = s_3 - 1/s_3, \qquad 2\gamma = s_3 + 1/s_3.$$

The invariants of the Neuberg curve are

$$I_2 = 3 + 2\rho\sigma, \qquad I_3 = -\gamma, \qquad I_4 = 9 + 12\rho\sigma + 4\rho^2.$$

It is evident that the covariant curve

$$C \equiv C_{32} + 2I_3C_{22} = 0 \tag{4.2}$$

cuts the Neuberg curve C_{12} at infinity which is one of the four special points under consideration. The same covariant must therefore pick out of C_{12} the remaining three points t_1, t_2, t_3 of this special set. That such is the case may be easily verified. However the covariant has eight points in common with C_{12} so that the original tetrad determines a second tetrad. Writing C_{12} and the covariant C as quadratics in $\bar{x}$ and calculating their eliminant we have the eight common points of C_{12} and C given as roots of the following equation of the eighth degree:

$$\begin{aligned} x^7 - 2\rho i x^6 + 2\rho i s_3 x^5 + (8\rho^2+7)s_3x^4 - 4s_3(6\rho i - \rho^2 s_3)x^3 \\ - 4s_3(5\rho i s_3 + 4\rho^2 - 2\rho^3 i s_3)x^2 - 8s_3^2(2\rho^2 + 1 - \rho^3 i s_3)x \\ + 4s_3^2(2\rho i + \rho^2 s_3) = 0, \end{aligned}$$

where the leading coefficient is zero since infinity is a root. Knowing t_1, t_2, t_3 to be roots, the above eliminant contains the factor

$$x^3 - 2\rho i x^2 - 2\rho i s_3 x - s_3.$$

The remaining factor

$$x^4 + 4\rho i s_3 x^2 + 8s_3 x - s_3(8\rho i + 4\rho^2 s_3)$$

has as roots the second tetrad. This tetrad consists of the incenter e_0 and the three centers of the escribed circles, e_1, e_2, e_3, of the triangle $t_1t_2t_3$. In terms of the t's we have

$$e_0 = \tau_1 - \tau_2 + \tau_3, \qquad e_1 = \tau_1 + \tau_2 - \tau_3,$$
$$e_2 = -\tau_1 - \tau_2 - \tau_3, \qquad e_3 = -\tau_1 + \tau_2 + \tau_3,$$

where $\tau_i = (s_3/t_i)^{1/2}$.

We object to (4. 2) on the ground that it is not homogeneous in the coefficients of C_{12}. An homogeneous relation is obtained as follows: It may easily be verified that the following non-homogeneous relation exists among the invariants of C_{12}, viz.

$$(4.3) \qquad I_4 I_3^2 - 6 I_2 I_3^2 + 9 I_3^2 - I_4 + I_2^2 = 0.$$

By inserting proper powers of $-2I_3C_{22}/C_{32}$ which is of degree two in the coefficients and by (4. 2) is equal to 1, the relation (4. 3) may be made homogeneous. We have then, after simplifying,

$$I_3^2\{I_4 C^3_{32} + 12 I_2 I_3 C^2_{32} C_{22} + 36 I_3^2 C_{32} C^2_{22} + 8 I_3 (I_4 - I_2^2) C^3_{22}\} = 0.$$

The part included in brackets is of degree 13 in the coefficients of C_{12} and of degree 6 in each of the variables. This covariant may cut out of C_{12} twenty-four points in three sets of two tetrads each, for it breaks up into three biquadratic factors of the type of (4. 2). Whether these three sets are all real, or one real and two complex is yet to be determined.

5. *Invariants of the bicubic.* In conclusion we record the results obtained by applying the method of Section 1 to the bicubic. We have the bicubic

C_{13}	1	$3x$	$3x^2$	x^3
1	a_{00}	a_{01}	a_{02}	a_{03}
$3\bar{x}$	a_{10}	a_{11}	a_{12}	a_{13}
$3\bar{x}^2$	a_{20}	a_{21}	a_{22}	a_{23}
$\bar{x}^3$	a_{30}	a_{31}	a_{32}	a_{33}

and its reciprocal form

C_{33}	1	x	x^2	x^3
1	A_{33}	A_{32}	A_{31}	A_{30}
$\bar{x}$	A_{23}	A_{22}	A_{21}	A_{20}
$\bar{x}^2$	A_{13}	A_{12}	A_{11}	A_{10}
$\bar{x}^3$	A_{03}	A_{02}	A_{01}	A_{00}.

The expansion of the determinant of the pencil

$$C_{13} + kC_{33},$$

as to powers of k, we find to be

$$I_4^3 k^4 + 2 I_4^2 I_2 k^3 + I_4 (18 I_4 + I_2^2) k^2 + 18 I_4 I_2 k + 81 I_4,$$

where the invariant I_4 is the determinant of the coefficients of C_{13} and

$$I_2 = (a_{00}a_{33} - a_{03}a_{30}) - 3(a_{01}a_{32} - a_{02}a_{31}) - 3(a_{10}a_{23} - a_{13}a_{20}) + 9(a_{11}a_{22} - a_{12}a_{21}).$$

THE JOHNS HOPKINS UNIVERSITY,
HAMILTON COLLEGE.

An Extension of Brahmagupta's Theorem.

By A. W. Richeson.

The question of cyclic polygons early attracted the attention of mathematicians. As early as 625 A. D. we find Brahmagupta, a Hindoo mathematician, working with the cyclic quadrangle.* About 1782 Lhuilier gave a formula for the radius of the circumcircle of an inscribed quadrangle in terms of the four sides, for both the convex and the crossed cases.† It is the purpose of this paper to study first the inscribed quadrangle and then to consider other polygons, and in particular the inscribed pentagon.

1. *Distances under Inversions.* Two points P and P' collinear with a given point O are said to be inverse points with respect to the point O if the product

$$\overline{OP} \cdot \overline{OP'} = K$$

where K is a real constant, O is the center, and K the constant of inversion. Given the inverse points P_0, P_1, $P_2 \cdots$ in the plane, let us denote by λ_{ij} the absolute distances of the points P_i and P_j. Under an inversion with respect to the point O as a center the distance λ_{12} becomes

$$K\lambda'_{12}/\lambda'_{01}\,\lambda'_{02}.$$

Accordingly a function of the distances which contains each suffix the same number of times merely acquires a factor under an inversion, that is to say, it is a relative invariant. Consequently for four points any one of the eight expressions

$$\lambda^2_{01}\,\lambda_{23}\,\lambda_{34}\,\lambda_{42} \pm \lambda^2_{02}\,\lambda_{34}\,\lambda_{41}\,\lambda_{13} \pm \lambda^2_{03}\,\lambda_{41}\,\lambda_{12}\,\lambda_{24} \pm \lambda^2_{04}\,\lambda_{12}\,\lambda_{23}\,\lambda_{31}$$

is a relative invariant under inversions.

Since equations of the type

$$\mu_1\overline{OP}_1{}^2 + \mu_2\overline{OP}_2{}^2 + \mu_3\overline{OP}_3{}^2 + \cdots + K = 0$$

are equations of circles, then

$$\lambda^2_{p1}\,\lambda_{23}\,\lambda_{34}\,\lambda_{42} \pm \lambda^2_{p2}\,\lambda_{34}\,\lambda_{41}\,\lambda_{13} \pm \lambda^2_{p3}\,\lambda_{41}\,\lambda_{12}\,\lambda_{24} \pm \lambda^2_{p4}\,\lambda_{12}\,\lambda_{23}\,\lambda_{31} = 0,$$

---

* Kaye, G. R., *Indian Mathematics*, Thacker, Spink Co., Calcutta (1915), p. [illegible].

† Archibald, R. C., "Area of the Quadrilateral," *Bulletin of the Mathematical Association of America*, Vol. 29, p. 32.

where p denotes the variable point and the λ's the absolute distances from the points, is the equation of a circle.

Attached to four points in the plane there is a system of eight circles obtained by considering the arrangement of the positive and the negative signs in

$$\sum^{4} \lambda^2_{p1} \lambda_{23} \lambda_{34} \lambda_{42} \pm = 0.$$

This system of circles may be classified as follows:

(1) When all the signs are positive, one nullipartite circle. This is the standard case.

(2) When two signs are positive and two negative, three unipartite circles, the Jacobian circles of the four points.

(3) When one sign is positive and three negative, four null or point circles.

In a paper in the *Proceedings of the London Mathematical Society,* R. Russell * considers seven of this system, the three unipartite or Jacobian circles and the four null or point circles. However, he failed to take into consideration the nullipartite circle.

I propose to determine the radii of this system of circles, and to consider what happens when the four points are on a circle.

2. *The Standard Case.* The equation of the circle in the standard case is

$$\lambda^2_{p1} \lambda_{23} \lambda_{34} \lambda_{42} + \lambda^2_{p2} \lambda_{34} \lambda_{41} \lambda_{13} + \lambda^2_{p3} \lambda_{41} \lambda_{12} \lambda_{24} + \lambda^2_{p4} \lambda_{12} \lambda_{23} \lambda_{31} = 0$$

where the λ's denote the absolute distances from the points and p is the variable point. In order to calculate the radius it is convenient to make a slight change in notation. Let x be the variable point p,

$$\begin{array}{ll} c = \lambda_{12} & e = \lambda_{24} \\ b = \lambda_{13} & f = \lambda_{34} \\ a = \lambda_{23} & d = \lambda_{41} \end{array}$$

Then the expression for the circle becomes

$$| x - x_1 |^2 aef + | x - x_2 |^2 bdf + | x - x_3 |^2 cde + | x - x_4 |^2 abc = 0$$

which may be written

$$(1) \quad (x - x_1)(\bar{x} - \bar{x}_1) aef + (x - x_2)(\bar{x} - \bar{x}_2) bdf + (x - x_3)(\bar{x} - \bar{x}_3) cde + (x - x_4)(\bar{x} - \bar{x}_4) abc = 0.$$

* Russell, R., "Geometry on the Quartic," *Proceedings of the London Mathematical Society,* Vol. 19 (1889), p. 56.

Expanding and collecting the terms of (1) we will be able to compute the center and radius of this circle in terms of the six distances a, b, c, d, e and f by equating the coefficients of (1) with those of the self conjugate equation of the circle

$$\alpha x\bar{x} + \bar{a}_1 x + a_1 \bar{x} + \beta = 0,$$

where the center $y = -a_1/\alpha$ and the radius

$$R^2 = (a_1\bar{a}_1 - \alpha\beta)/\alpha^2.$$

In this way we have the center

$$y = (-a_1)/\alpha = (aefx_1 + bfdx_2 + cdex_3 + abcx_4)/(aef + bfd + cde + abc)$$

which recalls the theorem that the centroid of four points is the sum of the products of the four points and their respective weights divided by the weights of the four points.

In a similar manner the radius is found to be

$$(2) \qquad \begin{aligned} R^2 &= (a_1\bar{a}_1 - \alpha\beta)/\alpha^2 \\ &= -2abcdef(ad + cf + be)/(aef + bfd + cde + abc)^2. \end{aligned}$$

This is the radius of the standard case of the circle and applies with change of sign to any one of the system of eight circles.

3. *The Jacobian Circle.* In order to obtain one of the Jacobian circles from

$$\sum^4 \lambda^2_{p1} \lambda_{23} \lambda_{34} \lambda_{42} \pm = 0$$

it is necessary to change the sign of one of the λ's, say λ_{13}, from positive to negative. We may then write

$$(3) \quad \lambda^2_{p1} \lambda_{23} \lambda_{34} \lambda_{42} - \lambda^2_{p2} \lambda_{34} \lambda_{41} \lambda_{13} + \lambda^2_{p3} \lambda_{41} \lambda_{12} \lambda_{24} - \lambda^2_{p4} \lambda_{12} \lambda_{23} \lambda_{31} = 0.$$

If this circle is on the point 1 then

$$-\lambda^2_{12} \lambda_{34} \lambda_{41} \lambda_{13} + \lambda^2_{13} \lambda_{41} \lambda_{12} \lambda_{24} - \lambda^2_{14} \lambda_{12} \lambda_{23} \lambda_{31} = 0.$$

This says that the point 1 is on the Jacobian circle if

$$-\lambda_{12}\lambda_{34} + \lambda_{13}\lambda_{24} - \lambda_{14}\lambda_{23} = 0.$$

Likewise for the points 2, 3 and 4 the same condition holds true. That is, the Jacobian circle is the circumcircle of the four points in the above order if $-\lambda_{12}\lambda_{34} + \lambda_{13}\lambda_{24} - \lambda_{14}\lambda_{23} = 0$. This is a known condition that four points be on a circle in the order 1, 2, 3, 4.

Since λ_{13} in

$$\sum^{4} \lambda_{p1}{}^2\lambda_{23}\lambda_{34}\lambda_{42} \pm = 0$$

was replaced by b and its sign changed to obtain the Jacobian circle, we must likewise change the sign of b in the expression (2) above for the radius of the standard case to obtain the radius of the Jacobian circle. Hence we may write

$$(4) \qquad R^2 = 2abcdef(ad + cf - be)/(aef - bfd + cde - abc)^2$$

for the radius of the Jacobian circle. But this value of the radius is indeterminate. For the condition that four points be on a circle is that $ad + cf - be = 0$ or that $aef - bfd + cde - abc = 0$. This is only one condition since the one implies the other.

In order to obtain the limit of this indeterminate expression (4) let us consider Cayley's relation connecting the mutual distances of any four points in the plane. In determinant* form it is as follows:

$$\begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & \lambda_{12}{}^2 & \lambda_{13}{}^2 & \lambda_{14}{}^2 \\ 1 & \lambda_{21}{}^2 & 0 & \lambda_{23}{}^2 & \lambda_{24}{}^2 \\ 1 & \lambda_{31}{}^2 & \lambda_{32}{}^2 & 0 & \lambda_{34}{}^2 \\ 1 & \lambda_{41}{}^2 & \lambda_{42}{}^2 & \lambda_{43}{}^2 & 0 \end{vmatrix} = 0.$$

Expanding and replacing the λ's by a, b, c, d, e and f we may write this in the form

$$\begin{aligned} a^2d^2(d^2 - e^2 - f^2) &+ a^2d^2(a^2 - b^2 - c^2) + b^2e^2(e^2 - d^2 - f^2) \\ &+ b^2e^2(b^2 - c^2 - a^2) + c^2f^2(f^2 - e^2 - d^2) + c^2f^2(c^2 - a^2 - b^2) \\ &+ a^2e^2f^2 + b^2f^2d^2 + c^2d^2e^2 + a^2b^2c^2 = 0. \end{aligned}$$

Now it is possible to express the left hand side in the form

$$\begin{aligned} (ad + cf + be)\{&2ad(a^2 + d^2) + 2be(b^2 + e^2) + 2cf(c^2 + f^2) \\ &-(ad + cf + be)\sum^{6} a^2\} + (aef + bfd + cde + abc)^2 \end{aligned}$$

which is of the form $g_3{}^2x + g_2y$ where $g_2 = ad + cf + be$ and $g_3 = aef + bfd + cde + abc$. From this we see that when $g_2 = 0$,

$$g_2/g_3{}^2 = -1/[2ad(a^2 + d^2) + 2be(b^2 + e^2) + 2cf(c^2 + f^2)]$$

and (2) becomes

$$R^2 = abcdef/\sum^{3} ad(a^2 + d^2).$$

Changing the sign of b, we have for the points 1, 2, 3, 4 on a circle (in this order)

---

* *Cayley's Collected Works*, Vol. I, p. 3; Salmon, Conic Sections, p. 134.

$$R^2 = -abcdef/[ad(a^2+d^2)-be(b^2+e^2)+cf(c^2+f^2)].$$

The two diagonals b and e may be expressed in terms of the other four distances * as

$$b^2 = (ad+cf)(af+cd)/(ac+df)$$

and

$$e^2 = (ad+cf)(ac+df)/(af+cd).$$

When these two relations are substituted in the expression for the radius above, it readily reduces to *

$$(4') \quad R^2 = (ac+df)(ad+cf)(af+cd)/(s-a)(s-c)(s-d)(s-f),$$

where $2s = a+c+d+f$. This gives the radius of the circumcircle of the four points in terms of the four distances a, c, d and f when the points are in the order x_1, x_2, x_3, x_4.

In order to obtain an expression for the radius of the Jacobian circle when the points are in the order x_1, x_2, x_4, x_3, that is when the quadrangle formed by the four distances a, c, d and f is crossed, it is necessary to change the sign of one of the sides, say f, in equation (4′). The radius then reduces to

$$R^2 = (ac-df)(ad-cf)(cd-af)/s(s-a-f)(s-c-f)(s-d-f),$$

where $2s = a+c+d+f$.

4. *The Inscribed Hexagon.* In the above sections the radius of the circumcircle of the quadrangle was calculated in terms of the four distances connecting the points for two cases, the convex quadrangle and the crossed quadrangle. In this and the following sections we shall consider other inscribed polygons, in particular the hexagon and the pentagon. For the hexagon we shall take the convex hexagon as the standard case.

We may think of the hexagon as made up of two quadrangles with a common side; i. e. the two quadrangles $abcx$ and $defx$ where x is the common side. From section (3) we have two equations for the circumcircle of the two quadrangles

$$(5) \quad R^2 = \frac{(ab+cx)(ac+bx)(ax+bc)}{(a+b+c-x)(b+c+x-a)(c+x+a-b)(x+a+b-c)}$$

and

$$(6) \quad R^2 = \frac{(de+fx)(df+ex)(dx+ef)}{(d+e+f-x)(e+f+x-d)(f+x+d-e)(x+d+e-f)}$$

These two equations must coexist if the hexagon is inscribed in the circle. We may further write the equality

* Hobson, *Plane Trigonometry* (Fifth Edition), p. 206.

$$\text{(A)}\qquad \frac{(ab+cx)(ac+bx)(ax+bc)}{(a+b+c-x)(b+c+x-a)(c+x+a-b)(x+a+b-}$$
$$=\frac{(de+fx)(df+ex)(dx+ef)}{(d+e+f-x)(e+f+x-d)(f+x+d-e)(x+d+e-}$$

This is an equation of the seventh degree in x and on eliminating x from equations (5) and (6) above we will obtain an equation of the seventh degree in R^2 of the type

$$\text{(7)}\qquad \lambda_0\alpha_0R^{14}+\lambda_1\alpha_1R^{12}+\lambda_2\alpha_2R^{10}+\lambda_3\alpha_3R^8+\lambda_4\alpha_4R^6 +\lambda_5\alpha_5R^4+\lambda_6\alpha_6R^2+\lambda_7\alpha_7=0$$

where the λ's are numerical constants and the α's are some function of the sides a, b, c, d, e, f. Simplifying the two equations and clearing of fractions we have

$$\text{(8)}\qquad R^2x^4+abcx^3+[a^2b^2+a^2c^2+b^2c^2-2R^2(a^2+b^2+c^2)]x^2 +abc(a^2b^2c^2-8R^2)x+[a^2b^2c^2-2R^2(a^2b^2+a^2c^2+b^2c^2) +R^2(d^4+e^4+f^4)]=0.$$

and

$$\text{(8}'\text{)}\qquad R^2x^4+defx^3+[d^2e^2+d^2f^2+e^2f^2-2R^2(d^2+e^2+f^2)]x^2 +def(d^2e^2f^2-8R^2)x+[d^2e^2f^2-2R^2(d^2e^2+d^2f^2+e^2f^2) +R^2(d^4+e^4+f^4)]=0.$$

If we set

$$\begin{array}{ll} \sigma_3=abc & s_3=def \\ \Sigma a^2b^2=a^2b^2+a^2c^2+b^2c^2 & \Sigma d^2e^2=d^2e^2+d^2f^2+e^2f^2 \\ \Sigma a^4=a^4+b^4+c^4 & \Sigma d^4=d^4+e^4+f^4 \end{array}$$

equations (8) and (8′) can be expressed in the condensed form

$$\text{(9)}\qquad R^2x^4+\sigma_3x^3+(\Sigma a^2b^2-2R^2\Sigma a^2)x^2 +(\sigma_3{}^3-8R^2\sigma_3)x+(\sigma_3{}^2-2R^2\Sigma a^2b^2+R^2\Sigma a^4)=0$$

and

$$\text{(9}'\text{)}\qquad R^2x^4+s_3x^3+(\Sigma d^2e^2-2R^2\Sigma d^2)x^2 +(s_3{}^3-8R^2s_3)x+(s_3{}^2-2R^2\Sigma d^2e^2+R^2\Sigma d^4)=0$$

and x can be eliminated from equations (9) and (9′) by Bezout's method of elimination.* This will give us a determinant of the fourth order and the seventh degree in R^2. We shall designate this determination by Δ_6:

* Salmon, *Higher Algebra*, p. 81.

$$
\Delta_6 = \left| \begin{array}{l}
(s_3 - \sigma_3), \\
(\Sigma d^2e^2 - 2R^2\Sigma d^2 - \Sigma a^2b^2 + 2R^2\Sigma a^2), \\
(s_3^3 - 8R^2 s_3 - \sigma_3^3 + 8R^2\sigma_3), \\
(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4 - \sigma_3^2 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4), \\
R^2[\Sigma d^2e^2 - 2R^2\Sigma d^2 - \Sigma a^2b^2 + 2R^2\Sigma a^2], \\
R^2[s_3^3 - 8R^2s_3 - \sigma_3^3 + 8R^2\sigma_3] \\
\qquad\qquad + \sigma_3\,[\Sigma d^2e^2 - 2R^2\Sigma d^2] - s_3[\Sigma a^2b^2 - 2R^2\Sigma a^2], \\
R^2(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^2 - \sigma_3^2 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4) \\
\qquad\qquad \cdot + \sigma_3(s_3^3 - 8R^2s_3) - s_3(\sigma_3^3 - 8R^2\sigma_3), \\
\sigma_3(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4) - s_3(\sigma_3^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4), \\
R^2(s_3^3 - 8R^2s_3 - \sigma_3^3 + 8R^2\sigma_3), \\
R^2(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4 - \sigma_3^2 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4) \\
\qquad\qquad + \sigma_3(s_3^3 - 8R^2s_3) - s_3(\sigma_3^3 - 8R^2\sigma_3), \\
\sigma_3(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4) - s_3(\sigma_3^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4) \\
\quad + (\Sigma a^2b^2 - 2R^2\Sigma a^2)(s_3^3 - 8R^2s_3) - (\Sigma d^2e^2 - 2R^2\Sigma d^2)(\sigma_3^3 - 8R^2s_3), \\
(\Sigma a^2b^2 - 2R^2\Sigma a^2)(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4) \\
\qquad\qquad - (\sigma_3^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4)(\Sigma d^2e^2 - 2R^2\Sigma d^2), \\
R^2(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4 - \sigma_3^2 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4) \\
\sigma_3(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4) - s_3(\sigma_3^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4) \\
(\Sigma a^2b^2 - 2R^2\Sigma a^2)(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4) \\
\qquad\qquad - (\sigma_3^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4)(\Sigma d^2e^2 - 2R^2\Sigma d^2) \\
(\sigma_3^3 - 8R^2\sigma_3)(s_3^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4) \\
\qquad\qquad - (s_3^3 - 8R^2s_3)(\sigma_3^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4)
\end{array} \right|
$$

Let us prove the following theorem:

THEOREM. *The product, $R_1^2 \cdot R_2^2 \cdot R_3^2 \cdot \cdot \cdot R_7^2$, of the seven roots of R^2 given by equation (7) after eliminating x from equations (9) and (9′) above is*

$$K \prod^{10} (abc - def) / \prod^{16} (a \pm b \pm c \pm d \pm e \pm f)$$

where K is a numerical constant and with the agreement that there shall be an odd number of negative signs in the factors of the denominator.

Let us examine the two coefficients $\lambda_0\alpha_0$ and $\lambda_7\alpha_0$ in equation (7) above. From equations (5) and (6) R^2 can become infinite when $x = a + b + c$ and $x = d + e + f$, that is, when at least one factor of the denominators of both (5) and (6) is zero. Hence R^2 is infinite if $a + b + c = d + e + f$ or $a + b + c - d - e - f = 0$. But since there are four factors in each denominator, for each factor of one there are four factors that go with it. This is the only way in which R^2 can become infinite since R^2 is not infinite when x is infinite. Hence $\lambda_7\alpha_0 = \lambda_0 \prod^{16} (a \pm b \pm c \pm d \pm e \pm f)$.

Now let us examine $\lambda_7\alpha_7$ in equation (7). From equations (5) and (6) R^2 can become zero if $ab+cx=0$ and $de+fx=0$. That is, R^2 will become zero if $abf-dec=0$; but since there are three factors in each numerator of (5) and (6), for each factor of the one there are three factors that go with it. Hence we have the product

$$\prod^{9}(abf-dec).$$

However on examining equation (7) in R^2, that is

$$\lambda_0\prod^{16}(a\pm b\pm c\pm d\pm e\pm f)R^{14}+\cdots+\lambda_7\alpha_7=0,$$

we find that each term should be of degree thirty, while $\prod^{9}(abf-dec)$ is of degree twenty-seven, hence there is a missing factor. If $x=\infty$, R^2 will be zero. That is $(abc-def)$ is zero and the missing factor is found, and we have $\lambda_7\alpha_7=\lambda_7\prod^{10}(abc-def)$; this is of degree thirty which is correct. Equation (7) may be written

$$\lambda_0\prod^{16}(a\pm b\pm c\pm d\pm e\pm f)R^{14}+\cdots+\lambda_7\prod^{10}(abc-def)=0.$$

The product of the seven roots of R^2 is given by $\lambda_7\alpha_7/\lambda_0\alpha_0$. Hence we have

$$R_1^2\cdot R_2^2\cdot R_3^2\cdots R_7^2=K\prod^{10}(abc-def)/\prod^{16}(a\pm b\pm c\pm d\pm e\pm f)$$

and the theorem is proved.

5. *The Inscribed Pentagon.* The same argument that was applied to the hexagon in the above section may be applied to the pentagon. Instead of two quadrangles with a common side inscribed in a circle we will have one quadrangle and a triangle having a common side to form the pentagon. That is the pentagon $abcde$ is formed by the quadrangle $abcx$ and the triangle dex where x is the common side.

The radius of the circumcircle of the pentagon is given by the equations

$$(10)\quad R^2=\frac{(ab+cx)(ac+bx)(ax+bc)}{(a+b+c-x)(b+c+x-a)(c+x+a-b)(x+a+b-}$$

and

$$(10')\quad R^2=\frac{d^2e^2x^2}{(d+e+x)(e+x-d)(x+d-e)(d+e-x)}$$

These two equations must coexist if the pentagon is inscribed in the circle and likewise an equality exist similar to (A) for the hexagon in the above section. This is an equation of the seventh degree in x and on eliminating x from equations (10) and (10′) we will obtain an equation of the seventh degree in R^2 of the type

$$\lambda_0\alpha_0 R^{14} + \lambda_1\alpha_1 R^{12} + \lambda_2\alpha_2 R^{10} + \lambda_3\alpha_3 R^8 + \cdots + \lambda_7\alpha_7 = 0\,, \tag{11}$$

where the λ's are numerical constants and the α's are some functions of the sides a, b, c, d, e.

The process of eliminating x from the two equations (10) and (10′) is similar to that of the hexagon but since the pentagon may be obtained directly from the hexagon by making one of its sides, say f, zero, it is not necessary to go through the elimination. On making $f = 0$ the determinant Δ_6 reduces to a determinant, which we shall designate by Δ_5 of the same order and degree in R^2 as Δ_6. The determinant Δ_5 is as follows:

$$\Delta_5 = \left| \begin{array}{l}
-\sigma_3, \\
(\Sigma d^2e^2 - 2R^2\Sigma d^2 - \Sigma a^2b^2 + 2R^2\Sigma a^2), \\
(-\sigma_3{}^3 + 8R^2\sigma_3), \\
(-\sigma_3{}^2 - 2R^2\Sigma d^2e^2 + R^2\Sigma d^4 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4), \\
R^2(\Sigma d^2e^2 - 2R^2\Sigma d^2 - \Sigma a^2b^2 + 2R^2\Sigma a^2), \\
R^2(-\sigma_3{}^2 + 8R^2\sigma_3) + \sigma_3(\Sigma d^2e^2 - 2R^2\Sigma d^2), \\
R^2(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4 - \sigma_3{}^2 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4), \\
\sigma_3(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4), \\
R^2(-\sigma_3{}^3 + 8R^2\sigma_3), \\
R^2(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4 - \sigma_3{}^2 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4), \\
\sigma_3(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4) - (\Sigma d^2e^2 - 2R^2\Sigma d^2)(\sigma_3{}^3 - 8R^2\sigma_3), \\
(\Sigma a^2b^2 - 2R^2\Sigma a^2)(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4) \\
\qquad\qquad - (\sigma_3{}^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4)(\Sigma d^2e^2 - 2R^2\Sigma d^2), \\
R^2(-R^2\Sigma d^2e^2 + R^2\Sigma d^4 - \sigma_3{}^2 + 2R^2\Sigma a^2b^2 - R^2\Sigma a^4) \\
\sigma_3(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4) \\
(\Sigma a^2b^2 - 2R^2\Sigma a^2)(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4) \\
\qquad\qquad - (\sigma_3{}^2 - 2R^2\Sigma a^2b^2 + R^2\Sigma a^4)(\Sigma d^2e^2 - 2R^2 d^2) \\
(\sigma_3{}^3 - 8R^2\sigma_3)(-2R^2\Sigma d^2e^2 + R^2\Sigma d^4)
\end{array} \right|$$

Let us prove a corresponding theorem for the pentagon as was proved for the hexagon.

THEOREM. *The product, $R_1{}^2 \cdot R_2{}^2 \cdot R_3{}^2 \cdots R_7{}^2$, of the seven roots of R^2, given by equations* (11) *after eliminating x from equations* (10) *and* (10′), *is*

[illegible]

[illegible]

From equations (10) and (10′) R^2 [illegible]

of (10) and (10′) are zero. Hence $a+b+c-d-e=0$. But since there are four factors in each denominator, for each factor of the one there are four factors that go with it. This is the only way in which R^2 can become infinite since R^2 is not infinite when x is infinite. Hence

$$\lambda_0\alpha_0=\lambda_0\prod^{16}(a\pm b\pm c\pm d\pm e).$$

Now let us examine $\lambda_7\alpha_7$. From equations (10) and (10′) R^2 can become zero if $ab+cx=0$ and either $d=0$ or $e=0$. That is to say $\lambda_7\alpha_7=\lambda_7(abcde)^\mu$, where μ must be determined by examining equation (11) in R^2. We find that each term is of degree thirty and μ must be 6. Equation (11) may be written

$$\lambda_0\prod^{16}(a\pm b\pm c\pm d\pm e)R^{14}+\cdots+\lambda_7(abcde)^6=0.$$

Since the product of the seven roots of R^2 is given by $\lambda_7\alpha_7/\lambda_0\alpha_0$ we have

$$R_1^2\cdot R_2^2\cdot R_3^2\cdots R_7^2=K(abcde)^6/\prod^{16}(a\pm b\pm c\pm d\pm e).$$

To determine the constant K we will consider the special case of the regular convex pentagon where $a=b=c=d=e$. The equations giving the radius of the circumcircle are

$$R^2=a^3(x+a)^3/(x+a)^3(3a-x)=a^3/(3a-x) \tag{12}$$

or $x=-a$ and

$$R^2=a^4x^2/(2a+x)(2a-x)x^2=a^4/(4a^2-x) \tag{12′}$$

or $x=0$. From (12) above

$$x=(3aR^2-a^3)/R^2.$$

Substituting this value of x in (12′) we obtain an equation of the second degree in R^2

$$5R^4-5a^2R^2+a^4=0.$$

$$\therefore\quad R^2=[5\pm(5)^{1/2}a^2]/10. \tag{13}$$

It should be noticed above that for $x=0$ in (12′) that

$$R^2=a^3/3 \tag{13′}$$

five times, since x was taken out of the two equations (12) and (12′) five times as a factor. Equation (13) gives two values of R^2 while (13′) gives five values which make up the total of seven values as shown by the general equation in R^2.

To determine the constant K required in the expression for the product of the seven roots of R^2 we have only to multiply these seven values together

$$R_1^2\cdot R_2^2\cdot R_3^2\cdots R_7^2=(1/5)(1/3)^5a^{14}.$$

Hence $K = (1/5)(1/3)^5$ and the expression for the product of the seven roots of R^2 is

$$(1/5)(1/3)^5 \ (abcde)^6 / \prod^{16} (a \pm b \pm c \pm d \pm e),$$

and the theorem is proved.

6. *The equation of the Seventh Degree in R^2.* To consider the general case of the pentagon where no two of the sides are equal we will center our attention on the determinant $\Delta_5 = 0$ obtained by eliminating x from equations (10) and (10′) in sec. 5 above. Upon expansion this determinant gives an equation of the type

$$\begin{aligned}(14) \qquad \lambda_0\alpha_0 R^{14} + \lambda_1\alpha_1 R^{12} + \lambda_2\alpha_2 R^{10} + \lambda_3\alpha_3 R^8 \\ + \lambda_4\alpha_4 R^6 + \lambda_5\alpha_5 R^4 + \lambda_6\alpha_6 R^2 + \lambda_7\alpha_7 = 0,\end{aligned}$$

where the λ's are constants and the α's are some function of the sides a, b, c, d, e. Instead of expanding $\Delta_5 = 0$ in full let us examine several of the coefficients of equation (14). We have already found in sec. 5 that

$$\lambda_0\alpha_0 = \lambda_0 \prod^{16} (a \pm b \pm c \pm d \pm e) \quad \text{and} \quad \lambda_7\alpha_7 = (abcde)^6.$$

The coefficients α_4, α_5, α_6 can be obtained in terms of the symmetric functions of the squares of the sides a, b, c, d, e of the pentagon by a partial expansion of Δ_5. Thus, if we set

$$S_1 = \sum^5 a^2,\ S_2 = \sum^{10} a^2b^2,\ S_3 = \sum^{10} a^2b^2c^2,\ S_4 = \sum^5 a^2b^2c^2d^2,\ S_5 = a^2b^2c^2d^2e^2,$$

these coefficients are as follows:

$$\begin{aligned}&\alpha_7 = S_5^3,\ \alpha_6 = S_5^2(2S_4 + 2S_1S_3 + 2S_2^2),\ \alpha_5 = S_5(4S_5S_2S_1 + S_4S_3S_1 - 4S_4^2),\\ &\alpha_4 = 2S_5^2S_2 + S_5S_4S_3 + S_3^4 + 2S_4^2S_1^4 - S_4^2S_2^2 - 4S_4^3 - 2S_3^2S_4S_1^2.\end{aligned}$$

Substituting these values for the α's we may write equation (14) as

$$\begin{aligned}&\prod^{16}(a \pm b \pm c \pm d \pm e)R^{14} + \cdots\\ &\quad + (2S_5^2S_2 + S_5S_4S_3 + S_3^4 + 2S_4^2S_1^4 - S_4^2S_2^2 - 4S_4^3 - 2S_3^2S_4S_1^2)R^6\\ &\quad + S_5(4S_5S_2S_1 + S_4S_3S_1 - 4S_4^2)R^4\\ &\quad + S_5^2(2S_4 + 2S_1S_3 + 2S_2^2)R^2 + S_5^3 = 0.\end{aligned}$$

If one of the sides, say e, of the pentagon is made zero then S_5 which is the product of the squares of the five sides becomes zero. Hence each of the coefficients α_7, α_6 and α_5 becomes zero and α_4, that is the coefficient of R^6, reduces to

$$\{ \prod^3 (a^2b^2 - c^2d^2) \}^2.$$

The remaining coefficients α_2 and α_3 are known in terms of S_1, S_2, S_3 and S_4 modulo S_5. The general equation of the pentagon reduces to

$$\prod^{16} (a \pm b \pm c \pm d) R^{14} + \cdots \{ \prod^{3} (a^2b^2 - c^2d^2) \}^2 R^6 = 0,$$

which is

$$[\{\prod^{4}(a \pm b \pm c \pm d)R^2 - \prod^{3}(ab + cd)\}\{\prod^{4}(a \pm b \pm c \pm d)R^2 - \prod^{3}(ab - cd)\}]$$

The first factor of this product when set equal to zero is the equation for the radius of the circumcircle of a convex quadrangle with sides a, b, c, d and the second factor when set equal to zero is likewise the equation for the radius of the circumcircle of a crossed quadrangle with the same sides a, b, c, d.

7. *The Seven Circumcircles.* We have already seen that the general equation for the pentagon obtained by eliminating x from equations (10) and (10′) is of the seventh degree in R^2. That is to say with five rods $a, b, c, d, e,$ which when jointed form a convex pentagon inscribed in a circle we may be able to obtain six other pentagons by forming all possible crossed pentagons with the same five rods. To obtain the seven pentagons and their circumcircles it is assumed that the five rods a, b, c, d, e are of lengths $a, a + \alpha, a + \beta, a + \gamma, a + \delta$ respectively where α, β, γ and δ are small compared with a and $\alpha \neq \beta \neq \gamma \neq \delta$. Let us consider some facts concerning these pentagons and their circumcircles. These circles are drawn with the sides of the pentagons in their order of length and we obtain the convex pentagon, two pentagons with one crossing similar to Fig. 1, three pentagons with two crossings similar to Fig. 2, and one with five crossings similar to Fig. 3. Of the seven circumcircles no two of them have equal radii. Further if we change the order of the sides we will obtain the same number of pentagons inscribed in the same circles with the same radii $R_1, R_2, \cdots, R_7$. However, there may be a difference in the number of pentagons with one and two crossings. It is also evident that the maximum number of crossings for the pentagon is five while it is impossible to obtain three or four crossings.

8. *The Inscribed* $(2n+1)$*-gon.* Let us consider the general case of the inscribed polygon of $(2n+1)$ sides. The $(2n+1)$-gon may be built up of two polygons of $(n+1)$ and $(n+2)$ sides having one side, say x, in common. Then two equations for the radius of the circumcircle of the type of (10) and (10′) of sec. 5 can be written which must coexist if the polygons are inscribed in a circle. Eliminating x the common side from these two equations we obtain an equation in R^2 of degree

$$\tfrac{1}{2}(2n+1)\binom{2n}{n} - 2^{2n-1}.*$$

* Dr. F. Morley, Lectures at the Johns Hopkins University, 1927-28.

This expression gives the number of polygons formed by $(2n+1)$ jointed rods when the $(2n+1)$ jointed rods for the convex case form an inscribed $(2n+1)$-gon. Equally well this expression gives the number of circles with radii $R_1, R_2, \cdots, R_{2n+1}$ in which these polygons are inscribed. Further it is easy to show that the series

$$\text{(L)} \qquad n+(n-1)\binom{2n+1}{1}+(n-2)\binom{2n+2}{2}\cdots$$

has exactly this expression for its sum.

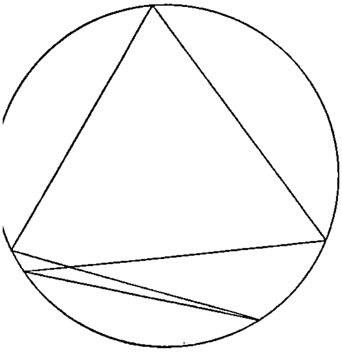


FIG. 1.

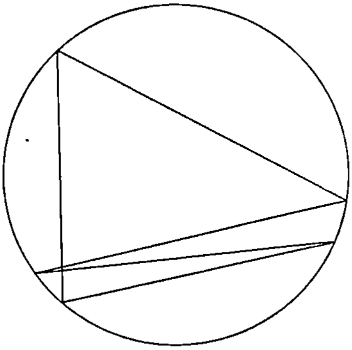


FIG. 2.

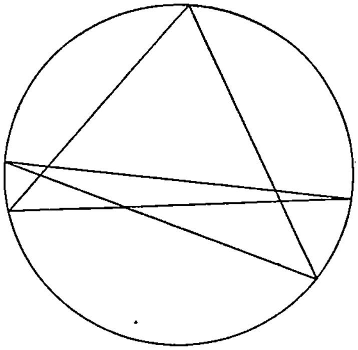


FIG. 3.

Let us examine a few special cases, say when $n=2, 3, 4$; i. e. when $2n+1=5, 7, 9$. Series (L) gives for these cases the following:

(M) $\qquad n=2, \quad 2n+1=5, \quad 2+5=7;$

(N) $\qquad n=3, \quad 2n+1=7, \quad 3+2(7)+21=38;$

(O) $\qquad n=4, \quad 2n+1=9, \quad 4+3(9)+2(36)+84=187.$

Expressions (M), (N) and (O) give the number of polygons formed by five, seven and nine jointed rods. Equally well these expressions give the number of circles in which these polygons are inscribed. In a similar manner we might extend this to $(2n+1)$ jointed rods.

Let us consider the numbers given by the last term in the left hand side of the expressions (M), (N) and (O), that is the numbers 5, 21 and 84. These are the number of times that each of the $(2n+1)$-gons degenerate into an equilateral triangle when all the sides are equal. In sec. 5 above we saw that in the case of the regular pentagon we obtained two inscribed pentagons, the regular convex and the regular star pentagon, and the degenerate case of the equilateral triangle five times over. For the heptagon the number of degenerate cases is 21, while for the 9-gon it is 84.

Let us examine each of these cases and see how this comes about. The case of the pentagon resolves itself into this: with five rods of equal length

labelled 1, 2, 3, 4, 5 forming a regular convex pentagon, how many distinct equilateral triangles can we form with these five rods? The answer is five. This may be accomplished as follows: bring sides 1 and 2 of the pentagon together until an angle of 60° is formed, these two sides are two of the sides, A and B, of the equilateral triangle. Now fold sides 3, 4 and 5 of the pentagon together forming the third side C of the triangle. This will give one of the triangles. Continue this procedure by taking two adjacent sides of the pentagon to form two sides of the triangle and in this manner we will form five distinct triangles as shown below.

Sides of triangle	A	B	C
Sides of pentagon	1	2	345
	2	3	451
	3	4	512
	4	5	123
	5	1	234

In the above we took the regular convex pentagon but the same results would have been obtained with the regular crossed or star pentagon.

For the heptagon we have the same problem. With seven rods of equal length labelled 1, 2, 3, 4, 5, 6, 7 how many distinct equilateral triangles can be formed? The answer is 21; this agrees with the number given by expression (N) above. They are as follows:

A	B	C
1	2	34567
.	.	.
7	1	23456
123	456	7
.	.	.
712	345	6
1234	56	7
.	.	.
7123	45	6

There are seven distinct triangles in each group which makes a total of twenty-one.

In a similar manner we can obtain eighty-four distinct equilateral triangles with nine labelled rods of equal length. We might extend this process on to polygons of $(2n+1)$ equal sides where we would obtain $\binom{2n+1}{n-1}$ equilateral triangles.

A Generalization of the Weddle Surface, of its Cremona Group, and of its Parametric Expression in Terms of Hyperelliptic Theta Functions.

By Arthur B. Coble.

---

1. Introduction. The two surfaces of greatest interest which occur in the study of the hyperelliptic theta functions of genus two are the two quartic surfaces discovered respectively by Weddle and by Kummer. The simplest definition of the Kummer surface with 16 nodes is transcendental—the coördinates of a point on the surface are proportional to theta functions of order two and like characteristic. The simplest definition of the Weddle surface with 6 nodes is projective—the surface is the locus of nodes of quadrics on six given points. These are the nodes of the locus. The most natural extension of the Kummer surface or 2-way is to the Kummer p-way, a variety, $K_p^{m(p)}$, of order $m(p)$ in a linear space, $S_{\nu(p)}$, of dimension $\nu(p)$, where

$$(1) \qquad \nu(p) = 2^p - 1, \qquad m(p) = p!\,2^{p-1}.$$

The coördinates of a point on K_p are proportional to theta functions of order two, of like characteristic, and of genus p (i. e. functions of p variables). This extension has been made by Wirtinger (for references cf.[1] pp. 94-109). On the other hand a natural extension of the Weddle surface is not immediately apparent, either in its projective definition or in such transcendental definitions as have been given.

The author has indicated[2] very briefly an extension which is based on the connection between a point of the Weddle surface and a projectively canonical form of the plane curve of genus two, the nodal planar quartic curve. It is the purpose of this article to elaborate this indication, and to study the generalized Weddle manifold W_p in the linear space S_{2p-1}, in part for a general value of p, but with greater particularity for $p = 3$. The canonical plane curve is the hyperelliptic curve H_p^{p+2} of order $p+2$, genus p, and with a p-fold point (cf.[1] pp. 125-135). A fundamental difference between K_p and W_p is to be noted. The manifold K_p depends upon the $p(p+1)/2$ moduli of the general theta function while W_p depends upon the $2p-1$ moduli of the hyperelliptic theta functions which are particular cases of the abelian theta functions defined by the algebraic curve of genus p with

$3p-3$ moduli. Whether there exists a type of manifold W_p for the more general manifold K_p is an open question. It is certain however that the peculiarities of the hyperelliptic K_p are more easily studied as a birational transform of the W_p we are about to discuss than as a particular case of the K_p with general moduli.

The Weddle surface is invariant under an abelian Cremona group G_{32} of type (1, 1, 1, 1, 1) which contains a particular element, or G_2, for which every point of W_2 is invariant. The factor group of G_{32} with respect to this invariant G_2 is a g_{16}. The permutations of the points of W_2 under this g_{16} are precisely those effected by adding the 16 half-periods to the arguments u_1, u_2 of a point of the surface when the coördinates of the point are expressed in terms of theta functions. In § 1 we discuss the extension of this group, an abelian $G_{2^{n+2}}$ of type $(1, 1, \cdots, 1)$ in a space S_n. For $n=2$ this group has been noted by Autonne ([3] p. 103); for $n=3$, by Kantor ([3] p. 213). The group is defined by $n+3$ points. *When n is odd,* it contains a particular element, or G_2, distinguished from the others by the fact that it involves all $n+3$ of the points symmetrically as F-points (fundamental or singular points) of the transformation. For $n>4$ this element, and the additional *regular transformation,*

$$x_i x_i' = 1 \qquad (i=0,1,\cdots,n), \tag{2}$$

are the only types of Cremona transformation defined symmetrically by a finite number of F-points ([4] II, p. 369). When $n=2p-1$ this element has a manifold of fixed points of dimension p *which we define to be* W_p. When W_p in S_{2p-1} is projected from one of the $2p+2$ F-points upon a W_p' in S_{2p-2}, the W_p' is invariant under the $G_{2^{n+2}}$ of this S_n $(n=2p-2)$. The latter group is simply isomorphic with the group of the addition of the 2^{2p} half-periods of the functions of genus p.

In § 2 a point of W_p is related to a type of planar hyperelliptic curve H_p^{p+2} with p-fold point at O and branch points of its unique linear series g_1^2 at $r_1, \cdots, r_{2p+2}$. Projective conditions on the branch points and O, expressed in terms of theta functions, are translated to W_p to obtain the parametric expressions for the coördinates of a point on W_p in terms of theta functions, analogous to those given by Schottky for the Weddle surface ([5]; [1] pp. 122-124).

The projective situation in S_{2p-1} then implies the existence of certain three-term relations connecting theta products. Such relations are immediate consequences of the general theory only in the case of small values of p. A systematic application of available three-term relations to geometric mani-

folds has been made by Schottky. These comprise products of genus two, and the Weddle surface ([1] pp. 122-124); products of genus three with u_1, u_2, u_3 subject to $\vartheta(u)=0$, and Cayley's dianodal surface ([1] pp. 181-184); modular products (i. e., products of even thetas for zero arguments) of genus three, and the planar quartic curve ([1] pp. 173-177); modular products of genus four when $\vartheta(0)=0$, and the space sextic of genus four on a quadric cone ([1] pp. 268-271); and modular products of genus four, and the ten nodes of a Cayley symmetroid ([1] pp. 273-276). To these cases treated by Schottky we are adding in § 3 the hyperelliptic types. A mapping of W_p upon the hyperelliptic Kummer manifold K_p is immediately apparent.

In § 4 the particular case, $p=3$, is developed in greater detail. Properties of the manifolds, W_3 in S_5, and W_3' in S_4, of orders 19 and 10 respectively, are derived. On W_3' in S_4 there appears a notable surface of order 11 with seven triple points, which is intimately related to the Weddle surface.

§ 1

THE CREMONA $G_{2^{n+2}}$ IN S_n.

2. The F-spaces of regular transformations in S_n. An F-point of a Cremona transformation of order m in S_n,

$$(3) \qquad T: \qquad x_i' = \phi_i(x) = (\phi_i x)^m \qquad (i=0,\cdots,n),$$

is a point f common to all the members, $\phi_i=0$, which define the homaloidal system. The correspondent of f under the transformation is indeterminate. The manifold common to the loci, $\phi_i=0$, must be made up of irreducible manifolds of various dimensions which are called the F-loci of T. These F-loci may meet in various loci of smaller dimension each of which is to be included among the F-loci of T. When we speak of a point on an F-locus of T it will be understood that the F-locus is irreducible and that the point is a generic point of it. No analysis of the nature and relative situation of these F-loci for transformations in higher space has been made and we pursue the matter only so far as is necessary for our immediate purpose.

Let f be any F-point, say a multiple point of order k on the generic member of the homaloidal system. It is a linear condition on the system that the members pass through f in such wise that a generic direction at f be on a member of the system. This linear condition on the system ϕ corresponds to a linear condition on the S_{n-1}'s, [illegible] $=0$, and is satisfied by all the $S_{[illegible]}$'s on a definite point x' which thus corresponds [illegible].

If y is a point on an S_{n-1} which is not on f and the direction at f is that of the line fy, the coördinates of the point x' are

$$(4) \qquad x_i' = (\phi_i f)^{m-k}(\phi_i y)^k.$$

Thus as y runs over the ∞^{n-1} points of $S_{n-1}(y)$ the point x' may run over a manifold P of dimension $n-1$. In this case f is an *F-point of the first kind,* and P is the *principal manifold* or *P-locus of the first kind* which corresponds to it. The number of such F-points is necessarily finite. Otherwise two generic lines in the space x' would correspond to two algebraic curves in the space x which would have different directions at an infinite number of common points. These F-points of the first kind are in general isolated from each other, though naturally, in particular cases, either coalescences may occur, or they may be located on one or more F-loci of higher dimension. The P-locus is the map of the $S_{n-1}(y)$ by the linear system

$$\sum_{i=0}^{i=n} \lambda_i (\phi_i f)^{m-k}(\phi_i y)^k = 0.$$

It may happen however that the mapping just mentioned is singular in that as y runs over $S_{n-1}(y)$ the point x' runs over a manifold M of dimension $n-1-r$. Then f is an *F-point of the* $(r+1)$*-th kind* and the directions about f correspond only to the points of M. If f is generic on an F-locus, F_r, of dimension r, then, as f runs over F_r, either the manifold M may run over a system ∞^r which makes up a P-locus of dimension $n-1$ *of the* $(r+1)$*-th kind,* or the manifold M may remain fixed and itself constitute an F_{n-r-1}-locus of the inverse transformation T^{-1}. It is this latter case only which occurs in regular transformations. We disregard the possibility that M may vary within the limits ∞^r and ∞^0, and prove that

(5) *The regular transformations T in S_n have, and are determined by, a finite number of F-points of the first kind. They have also a finite number of F-loci of the $(r+1)$-th kind $(r=1,\cdots,n-2)$, which are determined by the F-points of the first kind. An F_r-locus of the $(r+1)$-th kind of T is paired with an F_{n-r-1}-locus of the $(n-r)$-th kind of T^{-1} in such a way that the directions about each point of F_r correspond to all the points of F_{n-r-1} under T, and vice versa.*

A *regular transformation of order* n in S_n is by definition a transformation of the involutorial type (2) (when x, x' are in the same S_n) preceded or followed by a collineation. Hence the projective properties of this type can be deduced from the type (2). A *regular transformation* in S_n is a product of regular transformations of order n. In forming such regular pro-

ducts ST as we shall have occasion to use, the isolated F-points of T are placed in part at the isolated F-points of S^{-1} and, for the rest, in general position. The theory of the isolated F-points and the P-loci of the first kind of such regular transformations is developed in ([4] II §§ 4, 5). We first prove (5) for the simple type (2) in the form

$$(6) \qquad x_i' = \prod_{j=0}^{j=n} (x_j)/x_i \qquad (i = 0, 1, \cdots, n).$$

The members of the homaloidal system contain the $n + 1$ isolated F-points at the vertices of the reference $(n + 1)$-edron to the multiplicity $n - 1$, the $\binom{n+1}{2}$ F_1's joining any two of the vertices to the multiplicity $n - 2, \cdots$; in general the $\binom{n+1}{r+1}$ F_r's joining any $r + 1$ of the vertices to the multiplicity $n - r - 1$ $(r = 0, \cdots, n - 2)$. Let $x_0, x_1, \cdots, x_r, 0, 0, \cdots, 0$ be a generic F-point on one of these F_r's. A generic direction at this F-point on the line xy is determined by infinitesimal ϵ in

$$(7) \qquad x_0 + \epsilon y_0, \cdots, x_r + \epsilon y_r, \quad \epsilon y_{r+1}, \cdots, \epsilon y_n.$$

The transform of this by (6), after factoring out ϵ^{n-r-1}, then setting $\epsilon = 0$ and removing the further factor $x_0 x_1 \cdots x_r y_{r+1} \cdots y_n$, is

$$(8) \qquad 0, 0, \cdots, 0, 1/y_{r+1}, \cdots, 1/y_n.$$

Thus as y runs over any S_{n-1}, not on the F-point, the point, which corresponds to the direction from F on F_r to y, runs over the F_{n-r-1}-locus determined by the $n - r$ vertices complementary to the $r + 1$ which determine F_r, and this quite irrespective of the choice of F on F_r. Hence *opposite* F_r and F_{n-r-1} of the reference $(n + 1)$-edron correspond as in (5). In particular the point F_0 and opposite S_{n-1} are isolated F-point and corresponding P-locus of (2).

When (2) is preceded or followed by a collineation the resulting transformation T of order n has $n + 1$ isolated F-points which form one $(n + 1)$-edron, and their $(n + 1)$ P-loci form a dual $(n + 1)$-edron whose vertices, paired with those of the other, are the isolated F-points of T^{-1}. Then (5) holds as before except that opposite F_r and F_{n-r-1} are in different $(n + 1)$-edrons.

In a product of transformations T of order n the F-loci of the product must arise from F-loci of the particular factors. If a manifold M_r is converted, by applying the first k factors, into an F_r-locus of the $(k + 1)$-th [illegible] this factor converts it into an F_{n-r-1} which is in general transformed by further [illegible] into an [illegible]. Then M_r, M_{n-r-1} are F-loci paired as in (5). If, as factors are applied, M_{n-r-1} [illegible] M_r' then M_r, M_r' are ordinary images under the product. If however n is [illegible]

$n-r-1$, or $r=(n-1)/2$, a particular examination is necessary to see whether M_r, M_r' are F-loci. If they turn out to be F-loci they are paired as in (5). Hence (5) applies to all regular transformations.

We shall use from time to time two properties of the transformation (2) whose restatement for the general regular transformation *of order n* is obvious. These are:

(9) *A linear space S_{r+1} on F_r which meets the opposite F_{n-r-1} in y is transformed by* (2) *into a linear space S'_{r+1} on F_r which meets F_{n-r-1} in y'. The transformation* (2′) *between y, y' in F_{n-r-1} is of the type* (2) *and order $n-r-1$ with $(n-r)$-edron composed of the isolated F-points of* (2) *in F_{n-r-1}. The transformation between point x in S_{r+1} and point x' in S'_{r+1} is a regular transformation of order $r+1$ whose $r+2$ isolated F-points in S_{r+1} consist of the $r+1$ F-points of* (2) *in F_r and y; and whose $r+2$ isolated inverse F-points in S'_{r+1} consist of the $r+1$ F-points of* (2) *in F_r and y'.*

For, let an S_{r+1} on the F_r used in connection with (7) and (8) be defined by the equations $x_{r+1}/\lambda_{r+1}=\cdots=x_n/\lambda_n=k$. It meets the opposite F_{n-r-1} in the point $y=0,\cdots,0,\lambda_{r+1},\cdots,\lambda_n$. This S_{r+1} is transformed by (2) into the S'_{r+1} defined by $\lambda_{r+1}x_{r+1}=\cdots\lambda_n x_n=1/k$, which meets F_{n-r-1} in the point $y'=0,\cdots,0,1/\lambda_{r+1},\cdots,1/\lambda_n$, which proves the first property. From this point of view the transformation (2) in S_n has been called the *dilation* of the transformation (2′) in F_{n-r-1} ([6] pp. 256-258). When S_{r+1} is fixed by the choice of $\lambda_{r+1},\cdots,\lambda_n$, a point x on S_{r+1} has coördinates in S_n $x_0, x_1,\cdots,x_r, k\lambda_{r+1},\cdots,k\lambda_n$, and coördinates in S_{r+1}, $x_0, x_1,\cdots,x_r, k$. Its image in (2) has coördinates in S_n, $1/x_0, 1/x_1,\cdots,1/x_r, 1/k\lambda_{r+1},\cdots,1/k\lambda_n$, and coördinates in S'_{r+1}, $1/x_0, 1/x_1,\cdots,1/x_r, 1/k$, which proves the second property. Thus the transformation (2) in S_n appears as the product of subsidiary transformations of similar type in S_{n-r-1} and S_{r+1}.

3. The group $G_{2^{n+2}}$ defined by a set, P^n_{n+3} of $n+3$ points in S_n. Let $p_1, p_2,\cdots,p_{n+3}$ be the individual points of P^n_{n+3}. They are on a unique rational norm-curve, N^n, in S_n. If t is a parameter on N^n these points are determined by $t=t_1, t_2,\cdots,t_{n+3}$ respectively. Let I_{ij} $(i,j=1,\cdots,n+3)$ be the involutorial transformation of type (2) which interchanges p_i, p_j and which has isolated F-points at the remaining $n+1$ points of P^n_{n+3}. If these $n+1$ points are the reference $(n+1)$-edron and p_i, p_j are y, z respectively the equations of I_{ij} are

$$x_i'x_i=y_iz_i \qquad (i=0,\cdots,n). \tag{10}$$

The group generated by these $\binom{n+3}{2}$ involutions I_{ij} is characterized as follows:

(11) *The group generated by the involutions $I_{i_1i_2}$ contains 2^{n+2} involutorial elements $I_{i_1i_2\ldots i_{2l}}$ $(l = 0, 1, \cdots, (n+2)/2$ if n is even; $l = 0, 1, \cdots, (n+3)/2$ if n is odd) where $i_1, i_2, \cdots, i_{2l}$ are any $2l$ numbers selected from $1, \cdots, n+3$. This $G_{2^{n+2}}$ is abelian of type $(1, 1, \cdots, 1)$ and the law of multiplication is*

$$I_{i_1i_2\ldots i_{2l}} \cdot I_{j_1j_2\ldots j_{2m}} = I_{j_1j_2\ldots j_{2m}} \cdot I_{i_1i_2\ldots i_{2l}} = I_{k_1k_2\ldots k_{2n}},$$

where $k_1, \cdots, k_{2n}$ are the numbers found in one but not in both of the sets $i_1 \cdots i_{2l}$ and $j_1 \cdots j_{2m}$. The order of the subscripts in $I_{i_1\ldots i_{2l}}$ is not material.

Thus $I_{12}I_{13} = I_{13}I_{12} = I_{23}$, $I_{12}I_{34} = I_{13}I_{24} = I_{1234}$, etc. Given this law of multiplication, the elements are necessarily involutorial and the number of elements is $\binom{n+3}{0} + \binom{n+3}{2} + \cdots = 2^{n+2}$. In order to prove the law it is necessary to be able to define the individual elements. This may be done most conveniently by introducing the *principal* loci or P-loci of the set of points P^n_{n+3}. The P-loci of P^n_{n+3} are the P-loci which correspond to directions about the isolated F-points $p_1, \cdots, p_{n+3}$ under any regular transformation T whose isolated F-points are all included in the set P^n_{n+3}. Since T is generated by elements of type $I_{i_1i_2}$, all P-loci of the set P^n_{n+3} are obtained by applying these generating elements to the sets of directions at points p and by repeating the application to the new P-loci as they arise. The P-loci are all comprised in the aggregate,

$$\pi_{i_1\ldots i_{2k+1}} = (i_1^{k-1} \cdots i_{2k+1}^{k-1}\, i_{2k+2}^{k} \cdots i^k_{n+3})^k_{n-1} \qquad (1 \lesseqgtr 2k+1 \lesseqgtr n+3), \tag{12}$$

where the symbol on the right indicates a manifold of dimension $n-1$, and order k, with multiple points of order $k-1$ at $p_{i_1}, \cdots, p_{i_{2k+1}}$ and of order k at $p_{i_{2k+2}}, \cdots, p_{i_{n+3}}$. Included in the aggregate is

$$\pi_{i_1} = (i_1^{-1})^0_{n-1} \tag{13}$$

which represents the set of directions at p_i. That this P-locus, $\pi_{i_1\ldots i_{2k+1}}$, exists and is uniquely determined by the given order and multiplicities will be clear in the sequel where it appears that it can be transformed into the unique linear space on n given points. The rule for transforming these P-loci by the generating involutions is as follows:

(14) *The P-locus $\pi_{i_1\ldots i_{2k+1}}$ is transformed by the involution $I_{j_1j_2}$ into the P-locus $\pi_{i_1\ldots i_{2k+1}j_1j_2}$ provided that in the aggregate of subscripts, $i_1 \cdots i_{2k+1}j_1j_2$, like subscripts cancel.*

Because of the symmetry of π in (12) in the two complementary sets of indices it is sufficient to verify (14) for the three cases $I_{i_{2k}i_{2k+1}}$, $I_{i_{2k+1}i_{2k+2}}$, and $I_{i_{2k+2}i_{2k+3}}$. Under $I_{i_{2k}i_{2k+1}}$, π becomes

$$\left(i_1^{(n-1)k} \cdots i_{2k-1}^{(n-1)k}\; i_{2k}^{k-1}\; i_{2k+1}^{k-1}\; i_{2k+2}^{(n-1)k} \cdots i_{n+3}^{(n-1)k}\right)^{nk},$$

from which, due to the given multiplicities of π at the F-points of $I_{i_{2k}i_{2k+1}}$, the linear spaces $\pi_{i_j i_{2k} i_{2k+1}}$ for $j = 1, \cdots, 2k-1$ must each factor $k-1$ times, and the linear spaces $\pi_{i_j i_{2k} i_{2k+1}}$ for $j = 2k+2, \cdots, n+3$ must each factor k times. The order of the residual factor is $k - 1 = nk - (k-1)(2k-1) - k(n+3-2k-1)$; The multiplicity of the factor at $p_{i_1}, \cdots, p_{i_{2k-1}}$ is $(n-1)k - (k-1)(2k-2) - k(n-2k+2) = k - 2$; and the multiplicity at $p_{i_{2k+2}}, \cdots, p_{i_{n+3}}$ is $(n-1)k - (k-1)(2k-1) - k(n-2k+1) = k-1$. Hence the image of $\pi_{i_1 \ldots i_{2k+1}}$ is $\pi_{i_1 \ldots i_{2k-1}}$ or $\pi_{i_1 \ldots i_{2k+1} i_{2k} i_{2k+1}}$ according to the rule (14), provided there is a unique manifold of this type. The second and third cases can be verified similarly if in the second case account is taken also of the interchange of $p_{i_{2k+1}}$ and $p_{i_{2k+2}}$. Thus by a sequence of properly chosen involutions $I_{j_1 j_2}$, π can be transformed into $\pi_{i_1 i_2 i_3}$ which, as a linear space on $p_{i_4}, \cdots, p_{i_{n+3}}$, exists and is unique. Hence π exists and is unique.

A product of involutions $I_{j_1 j_2}$, or general member of the group (11), is defined when the P-loci corresponding to its isolated F-points, $p_1, \cdots, p_{n+3}$, are given; i. e., if the images of the P-loci π_{i_1} are given. Under the product $I_{12}I_{13}$ according to the rule of (14), π_1 becomes π_2 and then π_{123}; π_2 becomes π_1 and then π_3; π_3 becomes π_{123} and then π_2; while π_i $(i = 4, \cdots, n+3)$ becomes π_{i12} and then π_{i23}. Hence $I_{12}I_{13} = I_{13}I_{12} = I_{23}$. Under the product $I_{12}I_{34}$, π_i becomes π_{i12} and then π_{i1234}. This result, symmetric in the two sets of indices, $i = 1, 2, 3, 4$ and $i = 5, \cdots, n+3$, shows that $I_{12}I_{34} = I_{34}I_{12} = I_{13}I_{42} = \cdots = I_{1234}$. The continuation to complete the proof of (11) is obvious.

(15) *The general element $I_{i_1 i_2 \ldots 2l}$ of $G_{2^{n+2}}$ converts linear spaces into members of the homaloidal system*

$$\left(i_1^{(l-1)(n-1)} \cdots i_{2l}^{(l-1)(n-1)}\; i_{2l+1}^{l(n-1)} \cdots i_{n+3}^{l(n-1)}\right)^{1+l(n-1)}_{n-1}.$$

The P-locus of the isolated F-point p_{i_1} is $\pi_{i_2 \ldots i_{2l}}$; of the isolated F-point $p_{i_{2l+1}}$ is $\pi_{i_1 \ldots i_{2l} i_{2l+1}}$.

For, the theorem is true when $l = 0$ (the identity), and when $l = 1$ (a generating involution). Assuming that it is true for values of l up to and including $k-1$, we find by actual multiplication that $I_{i_1 \ldots i_{2k-2}} \cdot I_{i_{2k-1} i_{2k}}$ has a homaloidal system of the form (15) for $l = k$. If then $I_{i_1 \ldots i_{2l}}$ is

expressed as a product $I_{i_1 i_2} \cdot I_{i_3 i_4} \cdot \ldots \cdot I_{i_{2l-1} i_{2l}}$, the effect upon the sets of directions π_j is obtained from the rule (14).

Specific equations for the involution $I_{i_1 \ldots i_{2l}}$ in terms of the P-loci can be given. If we choose as reference points, $p_{i_3}, \cdots, p_{i_{2l}}, p_{i_{2l+1}}, \cdots, p_{i_{n+3}}$, the reference spaces opposite these points are respectively $\pi_{i_1 i_2 i_3}, \cdots, \pi_{i_1 i_2 i_{2l}}, \pi_{i_1 i_2 i_{2l+1}}, \cdots, \pi_{i_1 i_2 i_{n+3}}$. Then the image of $\pi_{i_1 i_2 i_3}$ under $I_{i_1 i_2 \ldots i_{2l}}$ is composed of the proper image of $\pi_{i_1 i_2 i_3}$ which, according to (14), is $\pi_{i_4 \ldots i_{2l}}$, and of the P-loci of the n isolated F-points on $\pi_{i_1 i_2 i_3}$. If the $(n+1)$ P-loci of the reference points are divided out, the equations of $I_{i_1 \ldots i_{2l}}$ take the form

$$
(16) \qquad
\begin{array}{ll}
\pi'_{i_1 i_2 i_3} = k_{i_3} \dfrac{\pi_{i_4 \ldots i_{2l}}}{\pi_{i_1 i_2 i_4 \ldots i_{2l}}}\,; & \pi'_{i_1 i_2 i_{2l+1}} = k_{i_{2l+1}} \dfrac{\pi_{i_3 \ldots i_{2l} i_{2l+1}}}{\pi_{i_1 i_2 i_3 \ldots i_{2l} i_{2l+1}}}\,; \\
\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\
\pi'_{i_1 i_2 i_{2l}} = k_{i_{2l}} \dfrac{\pi_{i_3 \ldots i_{2l-1}}}{\pi_{i_1 i_2 i_3 \ldots i_{2l-1}}}\,; & \pi'_{i_2 i_2 i_{n+3}} = k_{i_{n+3}} \dfrac{\pi_{i_3 \ldots i_{2l} i_{n+3}}}{\pi_{i_1 i_2 i_3 \ldots i_{2l} i_{n+3}}}\,.
\end{array}
$$

In order to obtain specific equations for the P-loci themselves we must observe that their geometric definitions as given in (12) in terms of the multiplicities at the points of P^n_{n+3} are too precise. For example, $\pi_{i_1 \ldots i_{2k+1}}$ of order k with k-fold points at $p_{i_{2k+2}}, \cdots, p_{i_{n+3}}$ has k-fold points at all the points of the linear S_{n+1-2k} on these given $(n+2-2k)$ k-fold points. Also if N^n is the unique rational norm curve of order n on P^n_{n+3}, then π has $(k-1)$-fold points not merely at $p_{i_1}, \cdots, p_{i_{2k+1}}$ but also at all points on N^n. This is a consequence of the following theorem in which the points $p_{i_1}, \cdots, p_{i_{2k+1}}$ do not appear explicitly.

(17) *The P-locus $\pi_{i_1 \ldots i_{2k+1}}$ $[0 < k \gtrless (n+2)/2]$ is the locus of the ∞^{k-1} linear spaces S_{n-k} which are $(n+1-k)$-secant to N^n, and on the fixed group of $n+2-2k$ points, $p_{i_{2k+2}}, \cdots, p_{i_{n+3}}$, which determines the k-fold locus S_{n+1-2k} of the P-locus.*

For, the spaces S_{n-k}, projected from S_{n+1-2k}, become spaces S_{k-2} in S_{2k-2} which are $(k-1)$-secant to the unique N^{2k-2} on the projections of $p_{i_1}, \cdots, p_{i_{2k+1}}$. According to the lemma (18) these spaces S_{k-2} make up the manifold of order k with $(k-1)$-fold points at the projections of $p_{i_1}, \cdots, p_{i_{2k+1}}$ and the proof of (17) is complete.

(18) *In the space S_{2k-2} the manifold of dimension $2k-3$ and order k with $(k-1)$-fold points at $p_1, \cdots, p_{2k+1}$ has $(k-1)$-fold points all along the N^{2k-2} on the P^{2k-2}_{2k+1}, and is the locus of the S_{k-2}'s which are $(k-1)$-secant to N^{2k-2}. If the coefficients of a binary $(2k-2)$-ic are taken as coördinates in S_{2k-2} and perfect powers represent points on N^{2k-2}, the equation of this manifold is given by the vanishing of the canonizant of the binary $(2k-2)$-ic.*

Let the binary $(2k-2)$-ic be represented symbolically by

$$(a_1 t)^{2k-2} = (a_2 t)^{2k-2} = (a_3 t)^{2k-2} = \cdots. \tag{19}$$

We use only one type of symbolic factor which, for two cogredient symbols, for two cogredient variables, and for contragredient symbol and variable, is to indicate respectively

$$(a_i a_j) = a_{i0} a_{j1} - a_{i1} a_{j0}; \qquad (t_i t_j) = t_{i0} t_{j1} - t_{i1} t_{j0}; \qquad (a_i t_j) = a_{i0} t_{j0} + a_{i1} t_{j1}. \tag{20}$$

The canonizant of the $(2k-2)$-ic is, symbolically,

$$(a_1 a_2)^2 (a_1 a_3)^2 (a_2 a_3)^2 \cdots (a_{k-1} a_k)^2; \tag{21}$$

and nonsymbolically,

$$\begin{vmatrix} a_0 & a_1 & \cdots & a_{k-1} \\ a_1 & a_2 & \cdots & a_k \\ \cdot & \cdot & \cdots & \cdot \\ a_{k-1} & a_k & \cdots & a_{2k-2} \end{vmatrix}. \tag{22}$$

The two-row minors of the determinant (22) of order k are of the form $a_i a_{i+j+k} - a_{i+j} a_{i+k}$, and vanish when the $(2k-2)$-ic is a perfect power. Hence the manifold (22) of order k has $(k-1)$-fold points on N^{2k-2}. The manifold of order k with $(k-1)$-fold points at $p_1, \cdots, p_{2k+1}$ is a P-locus in S_{2k-2} and is uniquely determined. It must therefore coincide with (22). If the canonizant vanishes, the $(2k-2)$-ic can be represented as a sum of $(k-1)$ perfect powers, and the point in S_{2k-2} represented by the $(2k-2)$-ic is in the S_{k-2} determined by the corresponding $k-1$ points on N^{2k-2}.

(23) *In the coördinate system in S_n with reference to N^n described in* (18)-(20), *the equation of the P-locus $\pi_{i_1 \ldots i_{2k+1}}$ is*

$$(a_1 a_2)^2 (a_1 a_3)^2 (a_2 a_3)^2 \cdots (a_{k-1} a_k)^2 (a_1 t_{i_{2k+2}}) \cdots (a_1 t_{i_{n+3}}) \cdots (a_k t_{i_{2k+2}}) \cdots (a_k t_{i_{n+3}}) = 0,$$

where $t_{i_{2k+2}}, \cdots, t_{i_{n+3}}$ are the parameters on N^n of $p_{i_{2k+2}}, \cdots, p_{i_{n+3}}$.

For, if, as in (17), a point of π is on an S_{n-k} which is $(n+1-k)$-secant to N^n, the corresponding binary n-ic can be expressed as a sum of $(n+1-k)$ perfect n-th powers of which $n+2-2k$ are the powers, $(t_{i_{2k+2}} t)^n, \cdots, (t_{i_{n+3}} t)^n$. The polar of $(t_{i_{2k+2}} t) \cdots (t_{i_{n+3}} t)$ as to this n-ic is a $(2k-2)$-ic which can be expressed as a sum of perfect $(2k-2)$-th powers of the remaining $k-1$ linear forms, whence the catalecticant of this polar vanishes.

4. Conjugate sets of F-loci of G_2^{n+2} of the j-th kind. The isolated

F-points and the P-loci of P^n_{n+3} may *be classed together as the F-loci of the first kind of* G_2^{n+2}. Of course a P-locus cannot be an actual F-locus of any element of G_2^{n+2}, but it is conjugate to the isolated F-points under G_2^{n+2} and for that reason may be classed with them.

We shall define the *k-th F-locus of* G_2^{n+2} *of the j-th kind*, $\pi^{(j)}{}_{i_1 i_2 \ldots i_{2k-j+2}}$ $[1 \leqq j \leqq (n+1)/2]$ to be, when $j \leqq k \leqq (n+j+1)/2$, the locus of dimension $n-j$ which is described by the ∞^{k-j} S_{n-k}'s on $p_{i_{2k-j+3}}, \cdots, p_{i_{n+3}}$ and on $k-j$ variable points of N^n; and when $(j-2)/2 \leqq k < j$, the locus of dimension $j-1$ which is described by the ∞^{k-j-1} S_k's on $p_{i_1}, \cdots, p_{i_{2k-j+2}}$, and on $j-k-1$ variable points of N^n. Of especial simplicity are the cases $k=j$ and $k=j-1$ for which the F-loci are linear spaces of respectively dimension $n-j$ on $p_{i_{j+3}}, \cdots, p_{i_{n+3}}$, and dimension $j-1$ on $p_{i_1}, \cdots, p_{i_j}$. These are corresponding F-loci under the element $I_{i_{j+1} i_{j+2}}$ of G_2^{n+2} (cf. **2**). Of interest also are the end cases in which the points of P^n_{n+3} occur symmetrically. These occur, for $k \geqq j$, when $2k-j = n+1$ and the F-locus consists of the ∞^{k-j} S_{n-k}'s which are $(k-j)$-secant to N^n, and, for $k < j$, when $2k = j-2$ and the F-locus consists of the $\infty^{j/2}$ $S_{(j-2)/2}$'s which are $(j/2)$-secant to N^n. Thus for $j=2$ we have the norm-curve N^n itself, i. e., $\pi^{(2)} = N^n$.

These F-loci are all of the same character as the P-loci, i. e. they are loci of multi-secant spaces of N^n. If we use the theorem of Castelnuovo (cf. [3] p. 155) which states that the spaces S_{m-1}, m-secant to N^n, make up an $M_{2m-1}^{\binom{n-m+1}{m}}$ on which N^n is a curve of multiplicity $\binom{n-m}{m-1}$, it is easy to show that

(24) *The F-locus* $\pi^{(j)}{}_{i_1 \ldots i_{2k-j+2}}$ $(k \geqq j)$, *has the order* $\binom{k}{j}$, *the multiplicity* $\binom{k}{j}$ *on the* S_{n-2k+j} *defined by* $p_{i_{2k-j+3}}, \cdots, p_{i_{n+3}}$, *and the multiplicity* $\binom{k-1}{j}$ *along* N^n. *The F-locus*, $\pi^{(j)}{}_{i_1 \ldots i_{2k-j+2}}$ $(k < j)$, *has the order* $\binom{n-k}{j-k-1}$, *the multiplicity* $\binom{n-k}{j-k-1}$ *on the* S_{2k-j+1} *defined by* $p_{i_1}, \cdots, p_{i_{2k-j+2}}$, *and the multiplicity* $\binom{n-k-1}{j-k-2}$ *along* N^n.

For, in the first case, the F-locus is made up of the ∞^{k-j} S_{n-k}'s on S_{n-2k+j} which are further $(k-j)$-secant to N^n. Projected from the S_{n-2k+j}, they become the ∞^{k-j} S_{k-j-1}'s which are $(k-j)$-secant to the projected N^{2k-j-1}. The latter locus has the order $\binom{k}{k-j}$ and contains N^{2k-j-1} with multiplicity $\binom{k-1}{k-j-1}$, which determine the order and multiplicities of the original locus. In the second case, the F-locus is made up of the ∞^{j-k-1} S_k's on S_{2k-j+1} which are further $(j-k-1)$-secant to N^n. Projected from S_{2k-j+1}, they become the ∞^{j-k-1} S_{j-k-2}'s which are $(j-k-1)$-secant to the

projected $N^{n-2k+j-2}$. The latter locus has the order $\binom{n-k}{j-k-1}$ and contains $N^{n-2k+j-2}$ to the multiplicity $\binom{n-k-1}{j-k-2}$.

Equations, analogous to those of (23), of linear systems of manifolds of dimension $n-1$, whose base is a particular F-locus, can be given. We observe that if a point of S_n lies in an S_{r+s-1} which is $(r+s)$-secant to N^n at $r+s$ points of which r are fixed at $p_1, \cdots, p_r$, then the binary n-ic, $(a_1t)^n$, whose coefficients determine the point, can be expressed as a sum of $r+s$ perfect n-th powers, r of which are $(t_1t)^n, \cdots, (t_rt)^n$. The polar $(n-r)$-ic, $(a_1t_1)\cdots(a_1t_r)(a_1t)^{n-r}$, can therefore be expressed as a sum of s perfect $(n-r)$-th powers. If then $(\alpha t)^{n-r-2s}$ is an arbitrary binary form of the order indicated, the polar $2s$-ic, $(a_1\alpha)^{n-r-2s}(a_1t_1)\cdots(a_1t_r)(a_1t)^{2s}$, can be expressed as a sum of s perfect $2s$-th powers, and its catalecticant vanishes for all values of the coefficients α. Naturally we must have $r+2s \lesseqgtr n$. The locus of such S_{r+s-1}'s is then defined by the linear system of manifolds of order $s+1$ [i. e., degree $s+1$ in the coefficients of $(a_1t)^n$],

$$(25) \qquad (a_1a_2)^2(a_1a_3)^2(a_2a_3)^2\cdots(a_sa_{s+1})^2(a_1\alpha_1)^{n-r-2s}\cdots(a_{s+1}\alpha_{s+1})^{n-r-2s}\cdot \\ (a_1t_1)\cdots(a_1t_r)\cdots(a_{s+1}t_1)\cdots(a_{s+1}t_r)=0,$$

i. e., the coefficients of the combinations of degree $s+1$ of the coefficients of $(\alpha t)^{n-r-2s}$ are themselves members of a linear system which has for its base the locus of the $(r+s)$-secant spaces. On applying this to the F-loci (24) we find that

(26) *For* $r=n-2k+j+1$, $s=k-j$, *and* $t_1, \cdots, t_r = t_{i_{2k-j+3}}, \cdots, t_{i_{n+3}}$ *the F-locus,* $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}}$ $(k \geqq j)$, *is defined by the linear system* (25) *in which the coefficients of* $(\alpha t)^{n-r-2s}=(\alpha t)^{j-1}$ *are the parameters; the F-locus* $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}}$ $(k<j)$, *is similarly defined when* $r=2k-j+2$, $s=j-k-1$, $t_1, \cdots, t_r = t_{i_1}, \cdots, t_{i_{2k-j+2}}$, *and* $(\alpha t)^{n-r-2s}=(\alpha t)^{n-j}$.

The number of F-loci of the j-th kind is either the number of even, or the number of odd, combinations of $n+3$ things; and thus is always 2^{n+2}, except when n is odd and $j=(n+1)/2$. In this case the 2^{n+2} types coalesce in pairs into 2^{n+1} types (see table at the end of this section). That they form a conjugate set under $G_{2^{n+2}}$ is a consequence of the theorem:

(27) *The F-locus,* $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}}$, *is transformed by* $I_{j_1\ldots j_{2l}}$ *into the F-locus,* $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}j_1\ldots j_{2l}}$ *where in the result like subscripts are to be deleted.*

It is sufficient to prove that the behavior of the F-loci is as stated in (27) for the generating involutions alone. We also have noted already that the linear F-loci for $k=j-1$ and $k=j$ are interchanged, i. e., that $\pi^{(j)}{}_{i_1\ldots i_j}$

is transformed by $I_{i_{j+1}i_{j+2}}$ into $\pi^{(j)}{}_{i_1\ldots i_{j+2}}$, and vice versa. We have to prove further only that, for $k \geqq j$, (a) $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}}$ is transformed by $I_{i_{2k-j+2}i_{2k-j+3}}$ into $\pi^{(j)}{}_{i_1\ldots i_{2k-j+1}i_{2k-j+3}}$; and that (b) $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}}$ is transformed by $I_{i_{2k-j+3}i_{2k-j+4}}$ into $\pi^{(j)}{}_{i_1\ldots i_{2k-j+4}}$; as well as that, for $k<j$, (c) $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}}$ is transformed by $I_{i_{2k-j+2}i_{2k-j+3}}$ into $\pi^{(j)}{}_{i_1\ldots i_{2k-j+1}i_{2k-j+3}}$; and that (d) $\pi^{(j)}{}_{i_1\ldots i_{2k-j+2}}$ is transformed by $I_{i_{2k-j+1}i_{2k-j+2}}$ into $\pi^{(j)}{}_{i_1\ldots i_{2k-j}}$.

In the case (a) both the original and transformed F-loci are made up of linear spaces on the $S_{n-2k+j-1}$ determined by the $n-2k+j$ points $p_{i_{2k-j+4}},\cdots,p_{i_{n+3}}$. According to (9) it is sufficient to prove the theorem for the projections of the F-loci from $S_{n-2k+j-1}$. Thus the proof for case (a) is reduced to the proof for case (a′): in S_{2k-j} to prove that the ∞^{k-j} S_{k-j}'s on $p_{i_{2k-j+3}}$ and $k-j$ variable points of N^{2k-j} is transformed by $I_{i_{2k-j+2}i_{2k-j+3}}$ into the ∞^{k-j} S_{k-j}'s on $p_{i_{2k-j+2}}$ and $k-j$ variable points of N^{2k-j}. Similarly case (b) is reduced to case (b′): in S_{2k-j+1} to prove that the ∞^{k-j} S_{k-j+1}'s on $p_{i_{2k-j+3}}, p_{i_{2k-j+4}}$ and $k-j$ variable points of N^{2k-j+1} is transformed by $I_{i_{2k-j+3}i_{2k-j+4}}$ into the ∞^{k-j+1} S_{k-j}'s on $k-j+1$ variable points of N^{2k-j+1}. Similar projections for cases (c) and (d) convert them into cases (c′) and (d′) which arise from cases (a′) and (b′) by setting $k=n-k$ and $j=n-j+1$. If now in case (a′) we set $2k-j=n$ and $k-j=r$, and in case (b′) we set $2k-j+1=n$ and $k-j+1=r$ they read as follows: Case (a″): in S_n to prove that the ∞^r S_r's on p_{n+3} and r variable points of N^n is transformed by $I_{n+2,n+3}$ into the ∞^r S_r's on p_{n+2} and r variable points of N^n; case (b″): in S_n to prove that the ∞^{r-1} S_r's on p_{n+2}, p_{n+3} and $r-1$ variable points of N^n is transformed by $I_{n+2,n+3}$ into the ∞^r S_{r-1}'s on r variable points of N^n. Thus the proof of (27) for the entire set of values of k is made to depend on the proof for the extreme value of k only and for the extreme generating involution $I_{n+2,n+3}$. The proof for case (a″) can be carried through precisely like the proof for case (b″) which we proceed to give.

In case (b″) let M_{2r-1} be the manifold described by the ∞^r r-secant S_{r-1}'s of N^n, and M'_{2r-1} that described by the ∞^{r-1} $(r+1)$-secant S_r's of N^n on p_{n+2}, p_{n+3}. The involution $I_{n+2,n+3}$ has isolated F-points at $p_1,\cdots,p_{n+1}$. Let g be any set of $n-2r$ of these $n+1$ F-points, say $p_1,\cdots,p_{n-2r}$, in an S_{n-2r-1}. If this S_{n-2r-1} be joined to each point of M_{2r-1}, and to each point of M'_{2r-1}, by an S_{n-2r}, the loci which thus arise are the two P-loci of, respectively, the ∞^r S_{n-r-1}'s on $p_1,\cdots,p_{n-2r}$ and r variable points of N^n, and the ∞^{r-1} S_{n-r}'s on $p_1,\cdots,p_{n-2r}$, p_{n+2}, p_{n+3} and $r-1$ variable points on N^n, i. e., the P-loci $\pi_{n-2r+1,\ldots,n+3}$ and $\pi_{n-2r+1,\ldots\,n+1}$. According to (14) these two P-loci are interchanged by $I_{n+2,n+3}$. If we prove that M_{2r-1} is the complete intersection of the various P-loci obtained from the $\binom{n+1}{n-2r}$ choices

of the set g, and that M'_{2r-1} is a similar complete intersection, then these two complete intersections are interchanged by $I_{n+2,n+3}$. We examine the case of M_{2r-1} first. Let x be any point of S_n on all the P-loci determined by M_{2r-1} and the choice of the set g, and let x be determined by the coefficients of the binary n-ic, $(at)^n$. Since x is on the particular P-locus given above, $(at)^n$ can be expressed in terms of the n-th powers

$$(t_1t)^n, \cdots, (t_{n-2r-1}t)^n, (t_{n-2r}t)^n, (s_1t)^n, \cdots, (s_rt)^n.$$

Thus the polar of $(t_1t)\cdots(t_{n-2r-1}t)$ as to $(at)^n$ is of order $2r+1$, and either t_{n-2r} is a root of the canonizant (of order $r+1$) of this polar, or the canonizant vanishes identically. But, as g varies, t_{n-2r} can be replaced by $t_{n-2r+1}, \cdots, t_{n+1}$, and the canonizant with these $2r+2$ distinct roots must vanish identically. But $(t_1t)\cdots(t_{n-2r-1}t)$ is any one of $\binom{n+1}{n-2r-1}$ products which may be formed from $(t_1t), \cdots, (t_{n+1}t)$ and any form of order $n-2r-1$ can be expressed linearly in terms of such products. If every polar of $(at)^n$ of order $2r+1$ can be expressed in terms of less than $r+1$ $(2r+1)$-th powers, then $(at)^n$ can be expressed in terms of less than $r+1$ n-th powers, and x is on M_{2r-1}. The proof that M'_{2r-1} is a complete intersection can be carried through in the same fashion and the proof for case (b″) is complete. The analogous proof for case (a″) completes the proof of (27).

The following tabular description of these F-loci in S_2, S_3, S_4, S_5 will be useful both by way of illustration and in connection with the more detailed discussion later of the manifolds W_3 in S_5 and W_3' in S_4:

j	*k*	*No.*	*Notation*	*Description*
		For P_5^2 in S_2:		
1	0	5	π_1	$^*p_1$;
1	1	10	π_{123}	$S_1(45)^1$;
1	2	1	π_{12345}	$M_1(1\cdots5)^2$.
		For P_6^3 in S_3:		
1	0	6	π_1	$^*p_1$;
1	1	20	π_{123}	$S_2(456)^1$;
1	2	6	$\pi_{1\ldots5}$	$M_2(1\cdots56^2)^2 = S_1$'s on p_6 to points of N^3;
2	0	1	$\pi^{(2)}$	N^3;
2	1	15	$\pi^{(2)}{}_{12}$	$S_1(12)^1$;
2	2	15	$\pi^{(2)}{}_{1\ldots4}$	$S_1(56)^1$;
2	3	1	$\pi^{(2)}{}_{1\ldots6}$	N^3;
		For P_7^4 in S_4:		
1	0	7	π_1	$^*p_1$;
1	1	35	π_{123}	$S_3(4567)^1$;
1	2	21	$\pi_{\ldots5}$	$M_3(1\cdots56^27^2)^2 = S_2$'s on p_6p_7 to points of N^4;

1	3	1	$\pi_{1\ldots 7}$	$M_3(1^2\cdots 7^2)^3 =$ bisecants of N^4;
2	0	1	$\pi^{(2)}$	N^4;
2	1	21	$\pi^{(2)}{}_{12}$	$S_1(12)^1$;
2	2	35	$\pi^{(2)}{}_{1\ldots 4}$	$S_2(567)^1$;
2	3	7	$\pi^{(2)}{}_{1\ldots 6}$	$M_2(1\cdots 67^3)^3 = S_1$'s on p_7 to points of N^4.

For P_8^5 in S_5:

1	0	8	π_1	$*p_1$;
1	1	56	π_{123}	$S_4(4\cdots 8)^1$;
1	2	56	$\pi_{1\ldots 5}$	$M_4(1\cdots 56^2 7^2 8^2)^2 = S_3$'s on $p_6p_7p_8$ to points of N^5;
1	3	8	$\pi_{1\ldots 7}$	$M_4(1^2\cdots 7^2 8^3)^3 = S_2$'s on p_8 to two points of N^5;
2	0	1	$\pi^{(2)}$	N^5;
2	1	28	$\pi^{(2)}{}_{12}$	$S_1(12)^1$;
2	2	70	$\pi^{(2)}{}_{1\ldots 4}$	$S_3(5678)^1$;
2	3	28	$\pi^{(2)}{}_{1\ldots 6}$	$M_3(1\cdots 67^3 8^3)^3 = S_2$'s on p_7p_8 to points of N^5;
2	4	1	$\pi^{(2)}{}_{1\ldots 8}$	$M_3(1^3\cdots 8^3)^6 =$ bisecants of N^5;
3	1	8	$\pi^{(3)}{}_1$	$M_2(1^4 2\cdots 8)^4 = S_1$'s on p_1 to points of N^5;
3	2	56	$\pi^{(3)}{}_{123}$	$S_2(123)^1$;
3	3	56	$\pi^{(3)}{}_{1\ldots 5}$	$S_2(678)^1$;
3	4	8	$\pi^{(3)}{}_{1\ldots 7}$	$M_2(1\cdots 78^4)^4 = S_1$'s on p_8 to points of N^5.

The doubling up of the types for $j = 2$ in S_3 and for $j = 3$ in S_5 is evident. This for odd spaces is an immediate consequence of the definition of the F-loci.

5. The symmetrical element in G_2^{n+2} when n is odd. If n is odd, say

(28) $$n + 3 = 2p + 2,$$

the group contains the symmetrical element $I_{1,2,\ldots,2p+2} = I$ whose homaloidal system (cf. 15),

(29) $$(1^{2p(p-1)}, \cdots, 2p + 2^{2p(p-1)})^{2p^2-1}_{2p-2},$$

contains all the points of P^{2p-1}_{2p+2} symmetrically. Under I the P-loci of the set of points are paired as follows:

(3) $$\pi_{i_1\ldots i_{2k+1}}, \quad \pi_{i_{2k+2}\ldots i_{2p+2}}.$$

Since I is invariant in the abelian group these pairs are permuted by G_2^{n+2} as entities.

According to the definition (12) of the P-loci, and according to (17), (18), the product of the members of a pair (30) is a spread of order p with $(p-1)$-fold points at P^{2p-1}_{2p+2} and with the $(p-2)$-fold curve N^{2p-1}. The pairs for which $k = 0$ are somewhat exceptional. The member π_{i_1}, the directions at p_{i_1}, is a spread of order 0 with multiplicity at p_{i_1} of virtual order -1. The complementary member, $\pi_{i_2\ldots i_{2p+2}}$, has at p_{i_1} a point of

order p, and has N^{2p-1} as a $(p-1)$-fold curve (cf. 18). Thus in the linear system Σ of spreads of order p with $(p-1)$-fold points P^{2p-1}_{2p+2}, the appearance of p_{i_1} to multiplicity p in $\pi_{i_2 \ldots i_{2p+2}}$ is the algebraic equivalent of the appearance of the complementary member π_{i_1}.

According to (18) the linear system Σ will contain a one-parameter system of spreads with $(p-1)$-fold points all along the curve N^{2p-1}. For, if t is any point of N^{2p-1}, the locus of the ∞^{p-1} S_{p-1}'s which are p-secant to N^{2p-1} and which pass through t has the order p and the $(p-1)$-fold curve N^{2p-1}. The equation of this one-parameter system with parameter t is, in the notation of $(19), \cdots, (23)$,

$$(31) \qquad (a_1a_2)^2(a_1a_3)^2(a_2a_3)^2 \cdots (a_{p-1}a_p)^2(a_1t)(a_2t)\cdots(a_pt) = 0.$$

When t takes the particular values $t_1, t_2, \cdots, t_{2p+2}$, the products (30) for $k=0$ and $i_1 = 1, \cdots, 2p+2$ are obtained.

The system Σ is an invariant linear system under the generators $I_{i_1i_2}$ and therefore an invariant linear system under $G_{2^{n+2}}$. The one-parameter system (31) is characterized by the fact that it contains the F-locus of the second kind, N^{2p-1}, to a multiplicity one greater than the normal multiplicity for spreads of Σ. But the system (31) also contains the F-locus of the second kind, $\pi^{(2)}{}_{1,2,\ldots,2p+2}$, the ∞^{p-1} S_{p-2}'s which are $(p-1)$-secant to N^{2p-1}, simply, i. e., to a multiplicity one greater than the normal multiplicity for spreads of Σ. Hence

(31a) *The linear system Σ of spreads of order p with $(p-1)$-fold points at P^{2p-1}_{2p+2} invariant under $G_{2^{n+2}}$ contains 2^{2p} members which break up into the members of a pair* (30). *The F-loci of the second kind of $G_{2^{n+2}}$ divide into 2^{2p} pairs, each a pair of conjugates under I,*

$$\pi^{(2)}{}_{i_1 \ldots i_{2k}}, \quad \pi^{(2)}{}_{i_{2k+1} \ldots i_{2p+2}}.$$

The members of Σ which contain one of these pairs of F-loci of the second kind to a multiplicity one greater than the normal multiplicity of Σ for the pair constitute a linear system of dimension p which contains a non-linear system of dimension one and degree p. Each of these non-linear systems contains $2p+2$ of the pairs (30) *whose parameters are projective to the fundamental $(2p+2)$-ic. Under I each member of Σ is invariant. Under the factor group of $G_{2^{n+2}}$ with respect to I, a $g_{2^{2p}}$, the 2^{2p} pairs* (30), *and the 2^{2p} non-linear systems, are permuted regularly.*

We prove in the next section that I has a locus, W_p, of fixed points. Anticipating this result for the moment we find the loci of fixed points of the various involutorial elements of $G_{2^{n+2}}$ for n odd or even.

6. The loci of fixed points of elements of G_2^{n+2}. We consider first the rather well-known cases, $n=2$, $n=3$. For $n=2$, the G_2^4 in S_2 has 10 elements of type I_{ij} and five of type I_{ijkl} $(i,\cdots,l=1,\cdots,5)$. The F-loci are all of P-type, the five sets of directions π_i, the ten lines $\pi_{ijk}=p_lp_m$, and the conic π_{12345}. The quadratic involution I_{ij} is non-perspective since p_i, p_j are not a line with any F-point. It has therefore four fixed points. The cubic involution I_{ijkl} has, however, a cubic locus of fixed points. For, under $I_{ijkl}=I_{ij}I_{kl}$ the line p_mp_i is transformed into p_mp_i whence every line of the pencil on p_m is invariant. The line p_ip_j is transformed into p_kp_l, whence the three degenerate conics, and therefore all conics, on $p_i,\cdots,p_l$ are invariant. The conics of this pencil are projected, each into itself, from the center p_m; and the locus of fixed points is generated by the pencil of conics, and the pencil of polar lines of p_m. Thus the locus is a cubic curve on P_5^2, tangent at p_m to π_{12345}, and on the diagonal triangle of $p_i,\cdots,p_l$.

For $n=3$ and G_2^5 with involutions I_{ij}, I_{ijkl}, $I_{1\ldots 6}$, the well known cubic involution I_{ij} has 8 fixed points which divide into two tetrahedra which, with the F-tetrahedron, make up a desmic system. The involution I_{1234} has a curve of fixed points. For, under $I_{1234}=I_{12}I_{34}$ the pencil of planes on p_5p_6 is invariant in such wise that the plane on p_i $(i=1,\cdots,4)$ is transformed into itself. Hence every plane E of the pencil is invariant, and in E I_{1234} is a quadratic transformation T with direct and inverse F-points at p_5, p_6 and with three corresponding pairs at the points where E cuts opposite edges of the tetrahedron, $p_1,\cdots,p_4$. This tetrahedron cuts E in a four-line and the pencil of conic envelopes on the four lines determines at any point of E an involution of pairs of tangents which contains the pairs of lines to the opposite points mentioned. The two involutions at points p_5 and p_6 of E construct the quadratic transformation T on E. One conic of the pencil touches the line p_5p_6 and the third F-point is the meet of the two further tangents to this conic form p_5, p_6. The four fixed points of T are the meets of the two tangents at p_5 to the two conics of the pencil on p_5 with the two tangents at p_6 to the two conics on p_6. Hence the locus of fixed points of I_{1234} is cut by a variable plane E on p_5p_6 in a variable set of four points.

Under I_{1234} the net of lines on p_6 is invariant. Projecting from p_6, the points $p_1,\cdots,p_5$ pass into a P_5^2 in S_2 and I_{1234} in S_2 determines the transformation I_{1234} on the net of lines on p_6. Hence under I_{1234} each line of a cubic cone with vertex at p_6 and on $p_1,\cdots,p_5$ is invariant. Similarly each line of a cubic cone with vertex at p_5 and on $p_1,\cdots,p_4,p_6$ is invariant. If x is on both cones it is transformed by I_{1234} both along the line p_6x, and along the line p_5x, and therefore is fixed. The curve of intersection of the

two cones, apart from the line p_5p_6, is an octavic curve with nodes at p_5, p_6 and on $p_1, \cdots, p_4$. This curve will project from either of its nodes into a doubly covered cubic curve.

Finally, as is well known, the involution $I_{123456} = I$ interchanges the planes π_{123} and π_{456}, and therefore leaves each member of the web of quadrics on P_6^3 invariant. Thus I is the involution which interchanges the seventh and eighth base points of a net of quadrics on six given points, and its locus of fixed points is the locus of nodes of quadrics of the web, i. e., the Weddle surface W_2.

The foregoing particular cases all illustrate the following theorem:

(32) *The element $I_{i_1 \ldots i_{2l}}$ of G_2^{n+2} in S_n has a locus of fixed points of dimension $l-1$, which, if $l > 2$, is projected from the linear space F_{n+2-2l} determined by $p_{i_{2l+1}}, \cdots, p_{i_{n+3}}$ into a Weddle manifold W_{l-1} in an S_{2l-3} which is covered 2^{n+3-2l} times.*

For, in the first place, if $l=1$, the regular transformation $I_{i_1i_2}$ of order n has 2^n fixed points which, for the canonical form (6), have coördinates $\pm 1, \pm 1, \cdots, \pm 1$. When n is odd these 2^n fixed points divide into two symmetrical sets of 2^{n-1} each according as the number of $+$ signs in the coördinates is even or odd.

In the general case let $I_{i_1 \ldots i_{2l}}$ be represented as the product $I_{i_1i_2}I_{i_3i_4} \cdots I_{i_{2l-1}\,2l}$. Each of these factors has F-points at $p_{i_{2l+1}}, \cdots, p_{i_{n+3}}$ in F_{n+2-2l}. Each factor, and therefore also $I_{i_1 \ldots i_{2l}}$, transforms the linear system of spaces E_{n+3-2l} on F_{n+2-2l} into itself [cf. (9)]. If the points $p_{i_1}, \cdots, p_{i_{2l}}$ are projected from F_{n+2-2l} into a set $q_{i_1}, \cdots, q_{i_{2l}}$ in an S_{2l-3}, the involution $I_{i_1 \ldots\ 2l}$ in S_{2l-3} has a locus of fixed points, W_{l-1}. Hence, under $I_{i_1 \ldots i_{2l}}$ in S_n, ∞^{l-1} of the spaces E are fixed. If x is a fixed point of $I_{i_1 \ldots i_{2l}}$, the space E on x must be fixed. Let E be a particular fixed space on F_{n+2-2l}, on which $I_{i_1 \ldots i_{2l}}$ effects the transformation T. This E is transformed by $I_{i_1i_2}$ into $E_{i_1i_2}$, $E_{i_1i_2}$ is transformed by $I_{i_3i_4}$ into $E_{i_3i_4}$, etc.; finally $E_{i_{2l-3}\,2l-2}$ is transformed by $I_{i_{2l-1}i_{2l}}$ back into E. According to (9) these constituent transformations of E' into E'' are regular and of order $n+3-2l$. Each has the fixed set, $p_{i_{2l+1}}, \cdots, p_{i_{n+3}}$, of $n+3-2l$ F-points. Hence the product T is a transformation of the same order $n+3-2l$ in E, which has a set of 2^{n+3-2l} fixed points, all of which project into the same point of W_{l-1}.

§ 2

A Transcendental Definition of W_p Suggested by the Planar Hyperelliptic Curve.

7. The manifold W_p and the planar hyperelliptic curve H_p. The theory of the hyperelliptic plane curve, H_p^{p+2}, of order $p+2$ and genus p with a p-fold point at O, under Jonquières transformation of the pencil of lines on O into itself is found in ([1] pp. 125-133). The curve possesses a g_1^2, cut out by the lines on O, whose branch points, $r_1, \cdots, r_{2p+2}$ are the contacts of tangents from O. We follow the exposition cited except that the branch point r_1, which figures [cf. [1] p. 131 (1)] as the fixed lower limit of the p normal integrals of the first kind, is here replaced by r_{2p+2}.

The planar set of points, $R^2{}_{2p+3}$, consisting of $r_1, \cdots, r_{2p+2}$ and $O = r_{2p+3}$, satisfies $p-1$ projective conditions [cf. [1] p. 127 (9)], and therefore depends upon $(4p+6)-8-(p-1) = 3p-1$ absolute projective constants. On the other hand, it is well known that two curves H_p, whose $2p+2$ branch lines on O are projectively equivalent, are themselves birationally equivalent; and thus H_p depends on only $2p-1$ absolute birational constants. Two curves, H_p, H_p' which are birationally equivalent, but not projectively equivalent, can be transformed into each other by a Cremona transformation of that particular type known as a Jonquières transformation J [cf. [1] p. 128 (12)]. If on H_p a p-ad of points be chosen, J can always be determined so that this p-ad is converted into the p-fold point O' of H_p'. As this p-ad varies on H_p the ∞^p curves H_p', all projectively distinct but birationally equivalent, are obtained. We denote the parameters of the $2p+2$ branch lines of H_p on O by $t_1, t_2, \cdots, t_{2p+2}$. These may be regarded as fixed for the ∞^p curves H_p'.

The planar set of points, $R^2{}_{2p+3}$, is "associated" with a set of $2p+3$ points P^{2p-1}_{2p+3} in a space S_{2p-1} in such wise that the points $p_1, \cdots, p_{2p+2}$, and $x = p_{2p+3}$, of P^{2p-1}_{2p+3} correspond respectively to the points $r_1, \cdots, r_{2p+2}$, and $O = r_{2p+3}$, of $R^2{}_{2p+3}$ (cf. [1] pp. 42-45). In S_{2p-1} the parameters of $p_1, \cdots, p_{2p+2}$ on the rational norm-curve, N^{2p-1}, determined by them, are projective to the lines of the pencil in S_2 from O to $r_1, \cdots, r_{2p+2}$ (cf. [1] p. 43 (c)). These parameters may therefore be taken to be $t_1, \cdots, t_{2p+2}$. Thus as H_p varies through the system of ∞^p curves H_p', the set P^{2p-1}_{2p+3} associated with $R'^2{}_{2p+3}$ has a set of $2p+2$ points fixed at $p_1, \cdots, p_{2p+2}$ and a variable point x which takes up ∞^p positions. We wish to prove that the locus of x is the manifold W_p, the locus of fixed points of the element $I = I_{1,2,\ldots,2p+2}$ of **5**.

The conditions on the planar set $R^2{}_{2p+3}$ that its points be the branch points

and multiple point of an H_p are the conditions that a Jonquières transformation J^{p+2} of order $p+2$ exist for which the set is congruent to itself in the identical order. In fact, if H_p exists, the J^{p+2} is precisely that Cremona involution which has H_p for a locus of fixed points. Moreover this J^{p+2} can be expressed as the product (cf. [1] p. 129),

$$(33) \qquad J^{p+2} = A_{O12}(12) \cdot A_{O34}(34) \ldots\ldots A_{O,2p+1,2p+2}(2p+1, 2p+2);$$

where it is understood that A_{O12} represents the quadratic involution with F-points at O, r_1, r_2, and with respective corresponding P-loci, r_1r_2, Or_2, Or_1; and that (12) represents the permutation of r_1, r_2. Thus $A_{O12}(12)$ represents the perspective quadratic involution for which every line on O is fixed.

If $R^2{}_{2p+3}$ is congruent to a set of points $R'^2{}_{2p+3}$ under A_{O12} (cf. [1] p. 10), its associated set P^{2p-1}_{2p+3} is congruent to a set of points P'^{2p-1}_{2p+3} under the regular transformation $A_{3,4,\ldots,2p+2}$ of order $2p-1$ in S_{2p-1} (cf. [1] p. 40 and p. 45 (8)) with F-points at $p_3, \cdots, p_{2p+2}$. Also the transposition (12) of r_1, r_2 entails the transposition of p_1, p_2 in S_{2p-1}. Hence to the transformation $A_{O12}(12)$ in S_2 there corresponds in S_{2p-1} the element I_{12} of the group $G_2{}^{n+2}$. If $R^2{}_{2p+3}$ is self-congruent in the identical order under the product (33), then P^{2p-1}_{2p+3} is self-congruent in the identical order under $I = I_{12}I_{34} \cdots I_{2p+1,2p+2}$; i. e., x of the set P^{2p-1}_{2p+3} is a fixed point of I. Conversely, if x is a fixed point of $I = I_{12} \cdots I_{2p+1,2p+2}$, the set P^{2p-1}_{2p+3} is self-congruent under I, and in S_2 the associated set $R^2{}_{2p+3}$ is self-congruent under the product (33). Hence $(J^{p+2})^2$ is a collineation which may be taken to be the identity, since only projective distinctions are being observed. Then J^{p+2} is involutorial, and has a curve H_p of fixed points for which $R^2{}_{2p+3}$ is the set of branch points and the multiple point. Hence

(34) *The locus of fixed points of the element I of $G_2{}^{n+2}$ $(n = 2p-1)$ is a manifold W_p of dimension p determined by the set P^{2p-1}_{2p+2} of points whose parameters on their norm-curve N^{2p-1} are $t_1, \cdots, t_{2p+2}$. As x varies on W_p the set $P^{2p-1}_{2p+3} = P^{2p-1}_{2p+2}$, x is associated with the sets $R'^2{}_{2p+3}$ of branch points and multiple point of the ∞^p projectively distinct hyperelliptic curves H_p' which have a fundamental $(2p+2)$-ic with roots projective to $t_1, \cdots, t_{2p+2}$.*

The $p-1$ projective conditions on the planar set $R^2{}_{2p+3}$ determined by H_p lead to geometric properties of the point x on W_p which are just sufficient to confine it to W_p. For example, if $p=2$, $r_1, \cdots, r_6$ are on a conic. But $r_1, \cdots, r_6$ are associated with the projection of $p_1, \cdots, p_6$ from x on the Weddle surface. Hence these projected points are likewise on a conic; i. e., x is the vertex of a quadric cone on $p_1, \cdots, p_6$ (cf. [1] p. 134). If

$p=3$, $r_1, \cdots, r_8$ and O are the base points of a pencil of cubics ([1] p. 127 (9)), and if three of the nine points are on a line, the complementary six are on a conic. Hence conics map the plane on a Veronese surface, V_2^4, and map the nine points, R_9^2, upon the associated P_9^5 in S_5. They also map the pencil of cubics on R_9^2 into a pencil of normal elliptic sextics, E^6, on P_9^5, which pencil describes V_2^4. Thus a point x of W_3 forms with P_8^5 the set of nine base points of ∞^1 E^6's; whereas a generic point of S_5 together with P_8^5 is on a unique E^6 (cf. [7] pp. 16-18). The V_2^4 is the locus of nodes of a cubic spread V_4^3; and thus there is isolated at x a particular member of the linear system Σ of (31a). The corresponding particular member, in the case $p=2$, is the nodal quadric on P_6^3 with node at x on W_2.

8. The loci on W_p which correspond to discriminant conditions on the planar set $R^2{}_{2p+3}$. If $R^2{}_{2p+3}$ is subjected to a projective condition, not dependent on the $p-1$ conditions which it already satisfies, the associated set P^{2p-1}_{2p+3} in S_{2p-1} is also subjected to an additional projective condition. This additional condition may fall, as we shall see, upon the points of P^{2p-1}_{2p+2} alone; but in general it will also involve the point x of W_p, and will, for given P^{2p-1}_{2p+2}, confine x to a locus on W_p of dimension $p-1$.

It is clear that those conditions on $R^2{}_{2p+3}$, which are most intimately related to the theory of the set under Cremona transformation, will be of greatest importance. Such are the "discriminant conditions" of the set $R^2{}_{2p+3}$ (cf. [1] p. 42). They are the conjugates under Cremona transformation of the condition that some two points of the set coincide. Nevertheless, in forming these conjugates, we use, not the general Cremona transformation with F-points in the set $R^2{}_{2p+3}$, but rather the Jonquières types with center at O and simple points in the set $r_1, \cdots, r_{2p+2}$. For it is only these latter types which convert an H_p into an H_p'. Thus we apply to these conditions only a sequence of quadratic transformations $A_{Or_ir_j}$.

Under these circumstances we find (cf. [1] p. 130) two conjugate sets of discriminant conditions. The first set arises from the coincidence of r_i and r_j, and is composed of $(p+1)(2p+1)$ pairs of conditions, $\delta_{ij}=0$ and $\delta(Or_ir_j)^1=0$, which indicate respectively that r_i and r_j coincide, and that O, r_i, r_j are on a line. Then either H_p acquires a node at $r_i=r_j$, or the line $(Or_ir_j)^1$ factors out from H_p; and the genus p is reduced.

The second conjugate set arises from a coincidence of r_i and O, which occurs when one of the p branches at O has a flex-point at O. The conjugates of this are of the form $\delta(O^kr_1\cdots r_{2k+3})^{k+1}=0$ $(k=-1, 0, \cdots, p-1)$ which represents the condition that there exists a curve of order $k+1$ with k-fold point at O and on $2k+3$, rather than only $2k+2$, of the branch

points; except that, for $k = -1$, it represents the coincidence of O and r_1. Then (cf. [1] p. 130 (17)).

(35) *The set $R^2{}_{2p+3}$ under Jonquières transformation has a conjugate set of 2^{2p+1} discriminant conditions which divide into 2^p pairs,*

$$\delta(O^k r_1 \cdots r_{2k+3})^{k+1} = 0, \quad \delta(O^{p-k-2} r_{2k+4} \cdots r_{2p+2})^{p-k-1} = 0$$
$$(k = -1, 0, \cdots, p-1),$$

either of which implies the other.

In fact, if one condition of a pair is satisfied, the existence of the Jonquières involution J^{p+2} defined by H_p implies that the other condition of the pair must also be satisfied.

We now translate these discriminant conditions on $R^2{}_{2p+3}$ to the corresponding discriminant conditions on the associated set P^{2p-1}_{2p+2} in S_{2p-1}. This is accomplished by observing that if two points of the one set coalesce, the corresponding two points of the associated set also coalesce; and that to the quadratic perspective element $A_{Or_1r_2}(12)$ in the plane there corresponds in S_{2p-1} the element I_{12} of the group $G_{2^{n+2}}$ (cf. [1] p. 45 (8), (9)). If then, in $R^2{}_{2p+3}$, r_1 and O coincide while, in P^{2p-1}_{2p+3}, p_1 and x coincide, we find, by applying the corresponding elements $A_{Or_2r_3}(23)$ and I_{23}, that $\delta(r_1r_2r_3)^1 = 0$ implies that x on W_p is on the P-locus π_{123}. Indeed the coincidence of x with p_1 means that x is on the P-locus π_1 and π_1 is transformed by I_{23} into π_{123} (cf. (14)). Continuing in this way we find that

(36) *The 2^p pairs of discriminant conditions* (35) *imply for the associated set P^{2p-1}_{2p+3} that x of W_p is on the section of W_p by either one of the pair of P-loci*

$$\pi_{1,2,\ldots,2k+3}, \quad \pi_{2k+4,\ldots,2p+2}.$$

That each P-locus of a pair cuts W_p in the same manifold is immediately evident from the fact that W_p is a locus of fixed points under I, and that the two P-loci are interchanged by I (cf. (31a)).

Reverting to the first type of discriminant condition on $R^2{}_{2p+3}$, we observe that $\delta_{ij} = 0$ implies in S_{2p-1} the coincidence of p_i, p_j; i. e., a double root of the fundamental $(2p+2)$-ic which determines the set P^{2p-1}_{2p+2} on the norm-curve N^{2p-1}. Also $\delta(Or_ir_j)^1 = 0$ implies that all of the $2p+2$ points p except p_i, p_j are on an S_{2p-2}. Hence N^{2p-1} must degenerate into an N^{2p-2} containing $2p$ of the points, and into the line p_ip_j which must meet N^{2p-2} in a point p. On N^{2p-2} the fundamental $(2p+2)$-ic consists of the $2p$ points and of p counted twice. These special cases of lower genus are not

considered further as we assume throughout that the fundamental $(2p+2)$-ic has distinct roots.

9. The equations in terms of theta functions of the sections of W_p by the P-loci of P^{2p-1}_{2p+2}. We recall some parts of the transcendental theory of H_p as given in ([1] pp. 131-133) with the single change that the common lower limit r_1 of the normal integrals of the first kind attached to H_p is here replaced by r_{2p+2}.

If $u = u_1, \cdots, u_p$ denote the sums of the respective normal integrals of the first kind with upper limits at a p-ad of points on H_p, then in general the p-ad determines $u_1, \cdots, u_p$ and conversely. The superposed p-ad, the correspondent of the given p-ad in the $g_1{}^2$ on H_p, then determines $-u$. The given p-ad and the superposed p-ad determine the same projective form of H_p, being interchanged by any Cremona involution (cf. [1] p. 126 (7)) which effects on H_p the involution whose pairs are in $g_1{}^2$.

The ∞^p projectively distinct curves H_p' which are birationally equivalent to H_p are distinguished by the choice of the p-ad of points on H_p which passes into the multiple point on H_p'. Thus the ∞^p points of W_p (cf. (34)) can be named by the sets of values $\pm u$, each determined by a p-ad of points on H_p (or by its superposed p-ad).

The P-curves of the planar set $R^2{}_{2p+3}$ which correspond under Jonquières transformation to directions at O are 2^{2p+1} in number, and divide into 2^{2p} pairs (cf. [1] p. 129 (15), (16)),

$$(37) \qquad L_k(r_1, \cdots, r_{2k+2}), \quad L_{p-k-1}(r_{2k+3}, \cdots, r_{2p+2}) \qquad (k = -1, 0, \cdots, p),$$

where L_k is a curve of order $k+1$ with a k-fold point at O and on $r_1, \cdots, r_{2k+2}$, and L_{-1} is the set of directions at O. The two curves of a pair cut out on H_p the same additional p-ad, whose parameters u' are related to u, the parameters of the p-ad at O, by the equations (cf. [1] p. 132 (7)):

$$(38) \qquad u' \equiv u + P_{1,2,\ldots,2k+2} \equiv u + P_{2k+3,\ldots,2p+2},$$

where $P_{1,2,\ldots,2j}$ is a half-period of the hyperelliptic theta functions in the basis notation. Since the p-ad at O is converted into the p-ad on L_k by the Jonquières transformation $A_{O12}(12) \cdots A_{O,2k+1,2k+2}(2k+1, 2k+2)$, the point x on W_p, which corresponds to the p-ad at O, is converted into the point x', which corresponds to the p-ad on L_k, by the element $I_{12} \cdots I_{2k+1,2k+2} = I_{1,2,\ldots,2k+2}$ of $G_2{}^{n+2}$. This element effects on W_p the same transformation as $I_{2k+3,\ldots,2p+2}$. Hence

(39) *If the points x of W_p are determined transcendentally by the value systems $\pm u = \pm u_1, \cdots, \pm u_p$, the factor group $g_{2^{2p}}$ of $G_{2^{n+2}}$ under I* (cf. (31a)) *effects on W_p the transformations expressed by the group* (38) *of additive half-periods.*

With the proper half periods represented as in (38) by the sets of $2k+2$, or the complementary set of $2(p-k)$ $(k=0,1,\cdots,p-1)$, indices selected from $2p+2$ indices, the 2^{2p} odd and even theta functions of the first order are represented in this basis notation by the sets of $p+1-2k$, or the complementary sets of $p+1+2k$, indices; i. e., they are

$$(40) \qquad \vartheta_{1,2,\ldots,p+1-2k}(u) = \vartheta_{p+2-2k,\ldots,2p+2}(u) \qquad [k=0,1,\cdots,(p+1)/2].$$

They are even or odd according as k is even or odd. If then p is odd there is a function $\vartheta(u) = \vartheta_{1,2,\ldots,2p+2}(u)$ which is even or odd according as $p \equiv 3$ or 1 (mod. 4). The hyperelliptic thetas differ from the general thetas in that all of the even functions (40), for which $k \neq 0$, vanish for the zero argument (cf. [8] pp. 456-464).

In ([1] p. 133) there is given a proof that the satisfaction of one of the 2^{2p} distinct discriminant conditions (35) entails the vanishing of a theta function for the arguments u of the p-ad at O on H_p. According to (36) these are also the conditions that x on W_p lie on a P-locus of the set P_{2p+2}^{2p-1}. We can therefore state at once that

(41) *If p is even, the condition that x on W_p with parameters u is on the P-locus $\pi_{1,2,\ldots,2k+3}$, or on $\pi_{2k+4,\ldots,2p+2}$ is*

$$\vartheta_{1,2,\ldots,2k+3}(u) = \vartheta_{2k+4,\ldots,2p+2}(u) = 0.$$

If p is odd, the condition that x is on the same pair of P-loci is

$$\vartheta_{2k+4,\ldots,2p+1}(u) = \vartheta_{1,2,\ldots,2k+3,2p+2}(u) = 0.$$

10. A provisional parametric representation of W_p. In giving a parametric representation of W_p by means of the hyperelliptic theta functions, it would suffice to give the section of W_p by $2p$ linear spaces which form a $2p$-edron in S_{2p-1}. But greater symmetry is obtained by giving the sections cut out by all the linear spaces on $2p-1$ points of the set P_{2p+2}^{2p-1}. These are the sections of W_p by the P-loci $\pi_{i_1 i_2 i_3}$ $(i_1, i_2, i_3 = 1, \cdots, 2p+2)$. According to (41) the parameters u of a generic point on this section constitute a zero of a definite theta function. But this particular linear section is characterized also by the fact that it passes through all the points of P_{2p+2}^{2p-1}.

except the points $p_{i_1}, p_{i_2}, p_{i_3}$. This behavior of the section is naturally indicated by the fact that its parametric representation contains factors which vanish when those discriminant conditions are satisfied, which imply a coincidence of x with a point p other than the three mentioned. The cases, p even, p odd, have somewhat different behavior.

When p is even, the directions at the points of P_{2p+2}^{2p-1} are those of π_{i_1}, and, according to (41), these are determined by $\vartheta_{i_1}(u)=0$. If then we set

$$(42) \qquad p \text{ even:} \qquad \pi_{i_1 i_2 i_3} = \vartheta_{i_1 i_2 i_3} \cdot \Pi / \vartheta_{i_1}\vartheta_{i_2}\vartheta_{i_3}, \quad \Pi = \vartheta_{i_1}\vartheta_{i_2} \cdots \vartheta_{i_{2p+2}},$$

we have a parametric representation of a p-way (u) in terms of theta products of order $2p$ and characteristic zero, which has the same behavior as W_p.

When p is odd, the index $2p+2$ is isolated. The directions at $p_1, \cdots, p_{2p+1}$ are determined respectively by $\vartheta_{1,2p+2}(u)=0, \cdots, \vartheta_{2p+1,2p+2}(u)=0$, while the directions at p_{2p+2} are determined by $\vartheta(u)=0$. The generic point on the section $\pi_{i_1, i_2, 2p+2}$ is determined by $\vartheta_{i_1 i_2}(u)=0$. The generic point on the section $\pi_{i_1 i_2 i_3}$ $(i_j \neq 2p+2)$ is determined by $\vartheta_{i_1, i_2, i_3, 2p+2}(u)=0$. The parametric representation is then

$$(43) \qquad p \text{ odd:} \qquad \begin{aligned} \pi_{i_1, i_2, 2p+2} &= \vartheta_{i_1 i_2} \cdot \Pi / \vartheta \vartheta_{i_1, 2p+2} \vartheta_{i_1, 2p+2}, \\ \pi_{i_1 i_2 i_3} &= \vartheta_{i_1 i_2 i_3, 2p+2} \cdot \Pi / \vartheta_{i_1, 2p+2} \vartheta_{i_2, 2p+2} \vartheta_{i_3, 2p+2}, \\ \Pi &= \vartheta \vartheta_{1,2p+2} \vartheta_{2,2p+2} \cdots \vartheta_{2p+1,2p+2}. \end{aligned}$$

Again the p-way is given in terms of functions of order $2p$ and characteristic zero.

In order to prove that the representations (42) and (43) are precisely those of the p-way, W_p, it is necessary to prove first that the functions to which the linear P-loci are equated satisfy the same linear relations as the P-loci. Secondly the coefficients of the linear relations satisfied by these functions must be such that the linear P-loci define a set of points projective to P_{2p+2}^{2p-1}. Thirdly the new functions which arise when u is replaced by $u+P_{12}$ (P_{12} a half period) must have a rational representation in terms of the original functions which is precisely that of the involution I_{12} of the Cremona $G_{2^{n+2}}$. The p-way will then necessarily be a locus of fixed points of the symmetrical element $I=_{1,2,\ldots,2p+}$ of $G_{2^{n+2}}$, and will coincide with W_p. These three requirements are disposed of in the following section.

§ 3

TRANSCENDENTAL DEFINITION OF W_p.

11. Three-term hyperelliptic theta relations. We divide the $2p+2$ indices of the basis notation into complementary sets of four, a, b, c, d; and of $2p-2$, $i_1 \cdots i_{2p-2}$. If to the argument of any one of the 2^{2p} odd and even theta functions of the first order we add the half period P_{abcd}, this function is converted, to within an outstanding exponential factor, into another one of the set, such that the product of the two is a theta function of second order and characteristic $abcd$. Of these 2^{2p-1} products half are even and half are odd. Of the $2^{2(p-1)}$ odd (or even) products only 2^{p-1} are linearly independent ([1] p. 56 (9)).

The odd products have the following typical forms:

$$p=2k: \quad \vartheta_{i_1 i_2 \ldots i_{2s-2} a} \vartheta_{i_1 i_2 \ldots i_{2s-2} bcd} \qquad (s=1, \cdots, k);$$
$$p=2k+1: \quad \vartheta_{i_1 i_2 \ldots i_{2s-1} a} \vartheta_{i_1 i_2 \ldots i_{2s-1} bcd}.$$

The theorem which we wish to prove reads as follows:

(44) *Any three of the following four products are linearly related:*

(a) p even: $\vartheta_{i_1}\vartheta_{i_2 i_3 i_4}$, $\vartheta_{i_2}\vartheta_{i_1 i_3 i_4}$, $\vartheta_{i_3}\vartheta_{i_1 i_2 i_4}$, $\vartheta_{i_4}\vartheta_{i_1 i_2 i_3}$;

(b) p odd: $\vartheta_{i_1 i_5}\vartheta_{i_2 i_3 i_4 i_5}$, $\vartheta_{i_2 i_5}\vartheta_{i_1 i_3 i_4 i_5}$, $\vartheta_{i_3 i_5}\vartheta_{i_1 i_2 i_4 i_5}$, $\vartheta_{i_4 i_5}\vartheta_{i_1 i_2 i_3 i_5}$;

(c) p odd: $\vartheta_{i_1 i_4}\vartheta_{i_2 i_3}$, $\vartheta_{i_2 i_4}\vartheta_{i_1 i_3}$, $\vartheta_{i_3 i_4}\vartheta_{i_1 i_2}$, $\vartheta\vartheta_{i_1 i_2 i_3 i_4}$

$$(i_1, \cdots, i_5 = 1, \cdots, 2p+2).$$

For, any $2^{p-1}+1$ of the $2^{2(p-1)}$ odd (or even) products must satisfy at least one linear relation, $R=0$, with constant coefficients which are not all zero. If the $2^{p-1}+1$ products can be arranged in a certain sequence, and if to each product, beginning with the first and continuing up to a certain point, there can be assigned a half period such that the product does not vanish for its assigned half period whereas the succeeding functions vanish for the half period, then there must exist a relation among the products to which such half periods have not been assigned. For, the substitution of the first half period in $R=0$ yields a zero in every term except the first, whence the coefficient of the first function is zero. Similarly by substituting the second half period in the simplified relation we find that the second function cannot appear, etc. We recall that the half period $P_{\alpha_1 \ldots \alpha_{2t}}$ substituted in the function $\vartheta_{\beta_1 \ldots \beta_j}(u)$ yields $E \cdot \vartheta_{\alpha_1 \ldots \alpha_{2t} \beta_1 \ldots \beta_j}(0)$, where E is a non-vanishing

exponential; and that $\vartheta_{\alpha_1\ldots\alpha_{2l}\beta_1\ldots\beta_j}(0)=0$ except when the indices $\alpha_1\cdots\alpha_{2l}\beta_1\cdots\beta_j$, with repeated indices deleted, number $p+1$.

When $p=2k$ we divide the $2p+2$ indices into three complementary sets $i_1,\cdots,i_{2k-1}$; $j_1,\cdots,j_{2k-1}$; a, b, c, d; and form the following sequence of products:

$$
\begin{array}{ll}
\vartheta_{i_1\ldots i_{2k-2}a}\vartheta_{i_1\ldots i_{2k-2}cbd}; & \vartheta_{i_1\ldots i_{2k-2}b}\vartheta_{i_1\ldots i_{2k-2}acd}; \\
\vartheta_{i_1\ldots i_{2k-4}a}\vartheta_{i_1\ldots i_{2k-4}bcd}; & \vartheta_{i_1\ldots i_{2k-4}b}\vartheta_{i_1\ldots {2k-4}acd}; \\
\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot & \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\
\vartheta_{i_1i_2a}\vartheta_{i_1i_2bcd}; & \vartheta_{i_1i_2b}\vartheta_{i_1i_2acd};
\end{array}
$$

$$\vartheta_a\vartheta_{bcd},\ \vartheta_b\vartheta_{acd},\ \vartheta_c\vartheta_{abd}.$$

It is to be understood that $\vartheta_{i_1\ldots i_{2k-2}a}\vartheta_{i_1\ldots i_{2k-2}bcd}$ represents the aggregate of products for the $\binom{2k-1}{2k-2}$ choices of $i_1,\cdots,i_{2k-2}$ from the set $i_1,\cdots,i_{2k-1}$; etc. The order of precedence in such an aggregate is not material. An aggregate on the right follows one on the left. The number of products is

$$2\left\{\binom{2k-1}{2k-2}+\binom{2k-1}{2k-4}+\cdots+\binom{2k-1}{2}\right\}+3$$
$$=(1+1)^{2k-1}-(1-1)^{2k-1}+1=2^{p-1}+1.$$

To this sequence of products we adjoin a sequence of half periods:

$$
\begin{array}{ll}
P_{i_{2k-1}b}; & P_{i_{2k-1}a}; \\
P_{i_{2k-3}i_{2k-2}i_{2k-1}b}; & P_{i_{2k-3}i_{2k-2}i_{2k-1}a}; \\
\cdot\quad\cdot\quad\cdot\quad\cdot & \cdot\quad\cdot\quad\cdot\quad\cdot \\
P_{i_3\ldots i_{2k-1}b}; & P_{i_3\ldots i_{2k-1}a},
\end{array}
$$

one for each product except the last three. An inspection of the two sequences shows that each half period is not a zero of its corresponding product and is a zero of each later product. Hence the last three products are linearly related which, for $a, b, c, d=i_1,\cdots,i_4$, proves (44a).

When $p=2k+1$ we divide the $2p+2$ indices into three complementary sets, $i_1,\cdots,i_{2k}$; $j_1,\cdots,j_{2k}$; and a, b, c, d. The following sequence of products and attached half periods serves the purpose:

$$
\begin{array}{lcl}
\vartheta_{i_1\ldots i_{2k-1}a}\vartheta_{i_1\ldots i_{2k-1}bcd} & : & P_{i_{2k}b}; \\
\vartheta_{i_1\ldots i_{2k-2}j_1b}\vartheta_{i_1\ldots i_{2k-2}j_1acd} & : & P_{i_{2k-1}i_{2k}j_1a}; \\
\vartheta_{i_1\ldots i_{2k-3}a}\vartheta_{i_1\ldots i_{2k-3}bcd} & : & P_{i_{2k-2}i_{2k-1}i_{2k}b}; \\
\vartheta_{i_1\ldots i_{2k-4}j_1b}\vartheta_{i_1\ldots i_{2k-4}j_1acd} & : & P_{i_{2k-3}\ldots i_{2k}j_1a}; \\
\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot & & \cdot\quad\cdot\quad\cdot\quad\cdot
\end{array}
$$

$$\begin{array}{lcl}
\vartheta_{i_1i_2j_1b}\vartheta_{i_1i_2j_1acd} & : & P_{i_3\ldots i_{2k}j_1a}; \\
\vartheta_{j_2\ldots j_{2k}b}\vartheta_{j_2\ldots j_{2k}acd} & : & P_{j_1a}; \\
\vartheta_{j_2a}\vartheta_{j_2bcd} & : & P_{j_1j_3\ldots j_{2k}b}; \\
\vartheta_{j_3a}\vartheta_{j_3bcd} & : & P_{j_1j_2j_4\ldots j_{2k}b}; \\
\cdot\quad\cdot\quad\cdot & & \cdot\quad\cdot\quad\cdot \\
\vartheta_{j_{2k}a}\vartheta_{j_{2k}bcd} & : & P_{j_1\ldots j_{2k-1}b};
\end{array}$$

$$\vartheta_{j_1a}\vartheta_{j_1bcd},\ \vartheta_{j_1b}\vartheta_{j_1acd},\ \vartheta_{j_1c}\vartheta_{j_1abd}.$$

Again the first products represent an aggregate determined by the choic the indices i, but the indices j, a, b, c, d are fixed as given. Thus the nun of products is

$$\binom{2k}{2k-1}+\binom{2k}{2k-2}+\cdots+\binom{2k}{2}+\binom{2k}{1}+3=(1+1)^{2k}+1=$$

The behavior of the products with respect to the attached half period the same as before whence there exists a linear relation among the last t which proves (44b).

If, in the relations (44b), u is replaced by $u+P_{i_1i_5}$, the relations (4 appear.

12. Transcendental definition of $Q^n{}_{n+3}$, w_p, and w_p'. We return consideration of the $\binom{2p+2}{3}$ theta products of order $2p$, $\pi_{i_1i_2i_3}$, which defined in (42), (43). There follows from (44) that

(45) (a) *Any three of the four functions,* $\pi_{i_2i_3i_4}$, $\pi_{i_1i_3i_4}$, $\pi_{i_1i_2i_4}$, $\pi_{i_1i_2i_3}$, *linearly related.* (b) *Of the* $\binom{k}{3}$ *functions* $\pi_{i_1i_2i_3}$ $(i_1, i_2, i_3=1,\cdots,k)$ *$k-2$ are linearly independent. A typical independent set is* $\pi_{i_1i_2i_3}$, π_{i} $\cdots$, $\pi_{i_1i_2i_k}$.

For, if p is even, (45a) is an immediate consequence of (44a). If odd, it is a consequence of (44b) or (44c) according as the index $2p+$ not, or is, included among $i_1,\cdots,i_4$. Then (45b) is true for $k=4$, we assume that it is true for values up to and including $k-1$. Given t $\pi_{i_1i_2i_k}$ in addition to $\pi_{i_1i_2i_j}$ $(j=3,\cdots,k-1)$, $\pi_{i_1i_ji_k}$, and similarly π_{i} $(h\neq j)$, can be obtained linearly. Hence $\pi_{i\,i_ji_k}$ can be obtained when π_{i} is added to the set $\pi_{i_1i_2i_3},\cdots,\pi_{i_1i_2i_{k-1}}$, and (45b) is true for the valu From this there follows that

(46) *Of the* $\binom{2p+2}{3}$ *functions* $\pi_{i_1i_2i_3}$ *only* $2p$ *are linearly independent. $2p$ independent functions are equated to the homogeneous coördinates point x in S_{2p-1}, then for variable u this point x ranges over a variety* w *S_{2p-1} of which the functions $\pi_{i_1i_2i_3}$ define linear sections. Of the* $\binom{2p+1}{3}$ *fı*

tions $\pi_{i_1i_2i_3}$ which omit a fixed index i_j $(j=1,\cdots,2p+2)$ only $2p-1$ are independent. They determine linear sections of w_p by spaces S_{2p-2} which are on a point q_{i_j}. The function $\pi_{i_1i_2i_3}$ defines the linear section of w_p by the S_{2p-2} on all the points q_{i_j} except q_{i_1}, q_{i_2}, q_{i_3}. Similarly $2p-1$ independent functions of the set $\pi_{i_1i_2i_3}$ which omit the index $2p+2$ define a variety w_p', and $2p+1$ points q'_{i_j} $(i_j=1,\cdots,2p+1)$, in S_{2p-2}, which is the projection of w_p and q_{i_j} in S_{2p-1} from the point q_{2p+2}.

Equations (42), (43) are parametric equations of the variety w_p. To obtain similar equations for w_p' we merely delete those linear spaces $\pi_{i_1i_2i_3}$ which contain the index $2p+2$. We have thus defined in S_{2p-1} a set of $2p+2$ points q_{i_j}, and in S_{2p-2} a set of $2p+1$ points q'_{i_j}; i. e., in S_n a set $Q^n{}_{n+3}$. For odd n the set is identified in **14** with the set of § 1 whose parameters $t_1,\cdots,t_{2p+2}$, on the N^{2p-1} containing the set, are projective to the roots of the fundamental $(2p+2)$-ic which defines the functions.

13. Transcendental equations of the P-loci of the set $Q^n{}_{n+3}$ of w_p and w_p'. We seek expressions, corresponding to the $\pi_{i_1i_2i_3}$ of (42) and (43), for the sections of w_p in S_{2p-1} by the entire system, $\pi_{i_1\ldots i_{2k+1}}$, of P-loci of Q^{2p-1}_{2p+2}. These are as follows:

(47) *When the coördinates x of a point of w_p are taken in such wise that the linear sections are given by (42) and (43), then the section of w_p by the P-locus $\pi_{i_1\ldots i_{2k+1}}$ of Q^{2p-1}_{2p+2}, of order k, is given by*

$$\begin{aligned}
p \text{ even: } & \pi_{i_1\ldots i_{2k+1}}=\vartheta_{i_1\ldots i_{2k+1}}\cdot\Pi^k/\vartheta_{i_1}\vartheta_{i_2}\cdots\vartheta_{i_{2k+1}};\\
p \text{ odd: } & \pi_{i_1\ldots i_{2k+1}}=\vartheta_{2p+2,i_1\ldots i_{2k+1}}\cdot\Pi^k/\vartheta_{i_1,2p+2}\cdots\vartheta_{i_{2k+1},2p+2};\\
& \pi_{2p+2,i_1\ldots i_{2k}}=\vartheta_{i_1\ldots i_{2k}}\cdot\Pi^k/\vartheta\vartheta_{i_1,2p+2}\cdots\vartheta_{i_{2k},2p+2}.
\end{aligned}$$

The proof is given here only for the more symmetric case when p is even. If in (44a) u is replaced by $u+P_{i_1i_5}$ the three functions,

$$\text{(a)}\qquad \vartheta_{i_5}\vartheta_{i_1\ldots i_5},\quad \vartheta_{i_1i_2i_5}\vartheta_{i_3i_4i_5},\quad \vartheta_{i_1i_3i_5}\vartheta_{i_2i_4i_5},$$

appear as linearly related. Then, multiplying by Π^2 and dividing by $\vartheta_{i_1}\cdots\vartheta_{i_4}\vartheta^2{}_{i_5}$, it appears that

$$\text{(b)}\qquad \vartheta_{i_1\ldots i_5}\cdot\Pi^2/\vartheta_{i_1}\cdots\vartheta_{i_5},\quad \pi_{i_1i_2i_5}\pi_{i_3i_4i_5},\quad \pi_{i_1i_3i_5}\pi_{i_2i_4i_5},$$

are linearly related. Hence the first term must represent a quadric with nodes at $q_{i_6},\cdots,q_{i_{2p+2}}$, and simple points at $q_{i_1},\cdots,q_{i_4}$. By virtue of the symmetry of this term in $i_1,\cdots,i_5$, and the consequent similar representations, this quadric must have a simple point at q_{i_5} also and must coincide with the P-locus $\pi_{i_1\ldots i_5}$. Thus we find the well known linear projective relation connecting the algebraic spreads,

(c) $$\pi_{i_1 \ldots i_5}, \quad \pi_{i_1 i_2 i_5} \pi_{i_3 i_4 i_5}, \quad \pi_{i_1 i_3 i_5} \pi_{i_2 i_4 i_5},$$

a relation which might have been obtained from the linear relation connecting $\pi_{i_1 i_2 i_3}$, $\pi_{i_1 i_2 i_4}$, $\pi_{i_1 i_3 i_4}$ by applying the involution $I_{i_4 i_5}$ defined by Q^{2p-1}_{2p+2}.

By adding $P_{i_6 \ldots i_{2k+1}}$ to u in (a) we find that

(d) $$\vartheta_{i_5 \ldots i_{2k+1}} \vartheta_{i_1 \ldots i_{2k+1}}, \quad \vartheta_{i_1 i_2 i_5 \ldots i_{2k+1}} \vartheta_{i_3 i_4 i_5 \ldots i_{2k+1}}, \quad \vartheta_{i_1 i_3 i_5 \ldots i_{2k+1}} \vartheta_{i_2 i_4 i_5 \ldots i_{2k+1}}$$

are linearly related. Multiplying by Π^{2k-2} and dividing by

$$\vartheta_{i_1} \cdots \vartheta_{i_4} \vartheta^2_{i_5} \cdots \vartheta^2_{i_{2p+2}},$$

the products,

(e) $$\pi_{i_5 \ldots i_{2k+1}} \cdot [\vartheta_{i_1 \ldots i_{2k+1}} \cdot \Pi^k / \vartheta_{i_1} \cdots \vartheta_{i_{2k+1}}], \quad \pi_{i_1 i_2 i_5 \ldots i_{2k+1}} \pi_{i_3 i_4 i_5 \ldots i_{2k+1}}, \quad \pi_{1 i_3 i_5 \ldots i_{2k+1}} \pi_{i_2 i_4 i_5 \ldots i_{2k+1}},$$

appear as related. On comparing this latter relation with the projective relation derived from (c) by applying the involution $I_{i_6 \ldots i_{2k+1}}$, which connects

(f) $$\pi_{i_5 \ldots i_{2k+1}} \pi_{i_1 \ldots i_{2k+1}}, \quad \pi_{i_1 i_2 i_5 \ldots i_{2k+1}} \pi_{i_3 i_4 i_5 \ldots i_{2k+1}}, \quad \pi_{i_1 i_3 i_5 \ldots i_{2k+1}} \pi_{i_2 i_4 i_5 \ldots i_{2k+1}},$$

we obtain the result (47) for p even. For, in these relations d and f, the coefficients of the last two terms must be such that, when in (d) the factor $\vartheta_{i_5 \ldots i_{2k+1}}$ appears in the first term, then in (f) the factor $\pi_{i_5 \ldots i_{2k+1}}$ must appear in the first term, provided (47) is assumed for smaller values of k. The residual factors, $\vartheta_{i_1 \ldots i_{2k+1}}$ and $\pi_{i_1 \ldots i_{2k+1}}$ are then uniquely determined and must also correspond.

The similar statement for w_p', the projection of w_p from q_{2p+2}, in S_{2p-2} with reference to the P-loci of Q'^{2p-2}_{2p+1}, the projection of Q^{2p-1}_{2p+2} from q_{2p+2}, is obtained by considering only the P-loci whose indices do not contain $2p + 2$.

14. Identification of the set Q^n_{n+3} with P^n_{n+3} of § 2. We have found in **12** that $\pi_{i_1 i_2 i_3}$ is the linear section of w_p by the S_{2p-2} on all of the points of Q^{2p-1}_{2p+2} except q_{i_1}, q_{i_2}, q_{i_3}. We wish to show that this set Q^{2p-1}_{2p+2} is projective to, and therefore, if we please, identical with the set P^{2p-1}_{2p+2} of § 2. It is necessary only to prove that the parameters t' of the points of Q on their norm-curve N'^{2p-1} are projective to the roots $t_1, \cdots, t_{2p+2}$ of the fundamental $(2p + 2)$-ic which defines the curve H_p^{p+2} of § 2 birationally, and which defines the theta functions used both in § 2 and § 3.

We denote by

(48) $$(i_1 i_2 \cdots i_k)$$

the determinant of the coördinates of an ordered set of k points in an S_{k-1}.

In particular, on an S_1 with non-homogeneous coördinate t, (i_1i_2) is $t_{i_1}-t_{i_2}$. For certain products of such differences we use the notation:

$$\begin{aligned} \{i_1i_2i_3\cdots\} &= (i_1i_2)(i_1i_3)(i_2i_3)\cdots, \\ \{i_1i_2\cdots;\ j_1j_2\cdots\} &= (i_1j_1)(i_2j_1)\cdots(i_1j_2)(i_2j_2)\cdots. \end{aligned} \tag{49}$$

The set P^{2p-1}_{2p+2} of § 1 with parameters $t_1,\cdots,t_{2p+2}$ on N^{2p-1} is associated with the set $P^1{}_{2p+2}$ of $2p+2$ points on S_1 with coördinates $t_1,\cdots,t_{2p+2}$; whence

$$(i_1i_2\cdots i_{2p}) = \lambda\cdot\epsilon_{i_1i_2\ldots i_{2p}i_{2p+1}i_{2p+2}}\cdot(i_{2p+1}i_{2p+2}), \tag{50}$$

where ϵ is $+1$ or -1 according as its indices are in an even or odd permutation from the natural order, and λ is a factor of proportionality.

If the reference system is so chosen that N^{2p-1} takes the canonical form

$$\begin{aligned} &x_i = t^{2p-1-i} \qquad (i=0,1,\cdots,2p-1), \qquad \text{then,} \\ &(i_1\cdots i_{2p}) = \{i_1\cdots i_{2p}\} = \lambda\cdot\epsilon_{i_1\ldots i_{2p+2}}\cdot(i_{2p+1}i_{2p+2}). \end{aligned} \tag{51}$$

All of the even hyperelliptic theta functions of the first order vanish when $u=0$ except the functions $\vartheta_{i_1\ldots i_{p+1}}=\vartheta_{i_{p+2}\ldots i_{2p+2}}$. Thomae[9] has proved that

$$\begin{aligned} c_{i_1\ldots i_{p+1}} = c_{i_{p+2}\ldots i_{2p+2}} &= \vartheta_{i_1\ldots i_{p+1}}(0) \\ &= \mu\cdot[\{i_1\cdots i_{p+1}\}\{i_{p+2}\cdots i_{2p+2}\}]^{1/4} \end{aligned} \tag{52}$$

where μ is a factor of proportionality (cf. also Zariski,[10] pp. 321 et seq.). These values enable us to determine the coefficients in the three-term relations (44) in terms of the differences of the roots $t_1,\cdots,t_{2p+2}$. In (44a),

$$\lambda_1\vartheta_{i_1}\vartheta_{i_2i_3i_4}+\lambda_2\vartheta_{i_2}\vartheta_{i_1i_3i_4}+\lambda_3\vartheta_{i_3}\vartheta_{i_1i_2i_4}\equiv 0,$$

let u be $P_{i_2i_3i_4j_1\ldots j_{p-1}}$, $P_{i_1i_3i_4j_1\ldots j_{p-1}}$, $P_{i_1i_2i_4j_1\ldots j_{p-1}}$ in turn. Then

$$\lambda_2c_{i_3i_4j_1\ldots j_{p-1}}c_{i_1i_2j_1\ldots j_{p-1}}+\lambda_3c_{i_2i_4j_1\ldots j_{p-1}}c_{i_1i_3j_1\ldots j_{p-1}}=0, \text{ etc.,}$$

where in either term a minus sign may occur (cf. [8] p. 240, VIII). But, according to (52),

(53)

$$\begin{aligned} c_{i_3i_4j_1\ldots j_{p-1}}c_{i_1i_2j_1\ldots j_{p-1}} = \eta\cdot[(i_1i_2)(i_3i_4)]^{1/2}\cdot[\{j_1\cdots j_{p-1}\}\{k_1\cdots k_{p-1}\}]^{1/2}\cdot \\ \cdot\{i_1\cdots i_4;\ j_1\cdots j_{p-1}k_1\cdots k_{p-1}\}^{1/4}. \end{aligned}$$

η being an undetermined fourth root of unity. Thus the identity is

$$\begin{aligned} &A_1\vartheta_{i_1}\vartheta_{i_2i_3i_4}+\eta A_2\vartheta_{i_2}\vartheta_{i_1i_3i_4}+\eta' A_3\vartheta_{i_3}\vartheta_{i_1i_2i_4}=0, \\ &A_r=[(i_si_t)(i_ri_4)]^{1/2} \qquad (r,s,t=1,2,3). \end{aligned} \tag{54a}$$

From 44(b) similar relations are derived:

(54b) $$A_1\vartheta_{i_1,2p+2}\vartheta_{i_2i_3i_4,2p+2}+\eta A_2\vartheta_{i_2,2p+2}\vartheta_{i_1i_3i_4,2p+2}+\eta' A_3\vartheta_{i_3,2p+2}\vartheta_{i_1i_2i_4,2p+2}\equiv 0;$$

(54b′) $$B_1\vartheta_{i_1,2p+2}\vartheta_{i_2i_3}+\eta B_2\vartheta_{i_2,2p+2}\vartheta_{i_1i_3}+\eta' B_3\vartheta_{i_3,2p+2}\vartheta_{i_1i_2}\equiv 0;$$

(54b″) $$B_1\vartheta_{i_2,2p+2}\vartheta_{i_1i_3}+\eta B_2\vartheta_{i_1,2p+2}\vartheta_{i_2i_3}+\eta' B_3\vartheta\vartheta_{i_1i_2i_3,2p+2}\equiv 0,$$

where B_r is obtained from A_r by interchanging i_4 and i_{2p+2}. Of these (b) is proved like (a); (b′) is obtained from (b) by adding $P_{i_4,2p+2}$ to u and then interchanging i_4 and $2p+2$; and (b'') is obtained from (b') by adding $P_{i_1i_2}$ to u.

Let π'_{ijk} be the theta function defined earlier with respect to the set Q^{2p-1}_{2p+2}; and set

$$G'_{ijk}=\pi'_{ijk}\cdot\{ijk\}^{1/2}.$$

Then all of the theta relations (54) take the typical form:

(55) $$(i_1i_4)G'_{i_2i_3i_4}+\eta(i_2i_4)G'_{i_3i_1i_4}+\eta'(i_3i_4)G'_{i_1i_2i_4}\equiv 0$$

where $G'_{i_1i_2i_3}$ is a linear space on all of the points of Q^{2p-1}_{2p+2} except q_{i_1}, q_{i_2}, q_{i_3}.

With respect to the points P^{2p-1}_{2p+2}, let $G_{i_1i_2i_3}$ be the linear space,

(56) $$G_{i_1i_2i_3}=\epsilon_{i_1\ldots i_{2p+2}}\cdot(i_4\cdots i_{2p+2}x),$$

on all of the points except p_{i_1}, p_{i_2}, p_{i_3}. The linear spaces, thus projectively defined, satisfy the projective relations,

$$(i_2i_3i_5\cdots i_{2p+2})(i_1i_5\cdots i_{2p+2}x)+(i_3i_1i_5\cdots i_{2p+2})(i_2i_5\cdots i_{2p+2}x)$$
$$+(i_1i_2i_5\cdots i_{2p+2})(i_3i_5\cdots i_{2p+2}x)\equiv 0.$$

These can be modified by using (56) and (50) to read:

(57) $$(i_1i_4)G_{i_2i_3i_4}+(i_2i_4)G_{i_3i_1i_4}+(i_3i_4)G_{i_1i_2i_4}\equiv 0.$$

Hence, according to (55) and (57), corresponding double ratios of four linear spaces on $Q^n{}_{n+3}$ and $P^n{}_{n+3}$ coincide to within fourth roots of unity; and the same is true of their respective associated sets, $t_1',\cdots,t'_{2p+2}$, and $t_1,\cdots,t_{2p+2}$, on S_1. But this can happen only if the fourth roots all are unity and the two linear sets are projective. For, if t_1', t_2', t_3' and t_1, t_2, t_3 are each projected into 0, 1, ∞ the coördinates of the remaining points are $t_i'=\eta_it_i$ $(i=4,\cdots,2p+2)$, where the η's are fourth roots of unity. Then also $(0,1;\ t_4'\,t_5')=\eta(0,1;\ t_4,t_5)$ which, for generic t_4, t_5, requires that $\eta=\eta_4=\eta_5=1$. Since the linear sets are projective, their associated sets, $Q^n{}_{n+3}$ and $P^n{}_{n+3}$, are also projective, and may be taken as superposed. Hence

(58) *The set of points $Q^n{}_{n+3}$ defined in* **12** *by the use of the functions* $\pi_{i_1i_2i_3}$

has parameters on the unique norm-curve N^n containing the set, which are projective to the roots $t_1, \cdots, t_{2p+2}$ $(n = 2p-1)$, or to the roots $t_1, \cdots, t_{2p+1}$ $(n = 2p-2)$, of the fundamental $(2p+2)$-ic which defines the functions. This set is projectively identical with the set P^n_{n+3} of §§ 1, 2.

We observe that the possible fourth roots of unity in the formulae (54) can be reduced, by the choice of the factor and the order of i, j, k in

$$G_{ijk} = \dot{\pi}_{ijk} \cdot \{ijk\}^{1/2}, \tag{59}$$

to the square roots ± 1 which occur in the projective relations (57). Unfortunately the basis notation is not sufficiently precise to distinguish between these signs (cf. [1] p. 86). But this difficulty can be removed by a transition, as in (59), to the projective form in which the signs of the expressions used depend only on the order of the points concerned.

15. Identification of w_p, w_p' with the W_p, W_p' of § 2. Let the set of points defined by the linear spaces $\pi_{i_1 i_2 i_3}$ in (42), (43) be denoted now by P^n_{n+3}. If they are expressed in terms of a coördinate x in S_n, then (42), (43) can be solved for x to obtain explicit expressions for these coördinates as theta functions of u of order $2p$. As u varies, the point x, thus determined by u, runs over w_p. We consider the relation between the two points x, x' on w_p determined by u and $u' \equiv u + P_{i_1 i_2}$. The point x' is determined by the linear expressions, $\pi'_{i_1 i_2 i_3}$, which arise when u in (42), (43) is replaced by $u + P_{i_1 i_2}$. Hence, in the case p even, which is the only one we consider in detail, typical values of the $\pi'_{i_1 i_2 i_3}$ are:

$$\begin{array}{lll}
\text{(a)} & \rho\pi'_{i_1 i_2 i_3} = \pm\, \vartheta_{i_3} & \cdot\ \Pi'/\vartheta_{i_1}\vartheta_{i_2}\vartheta_{i_1 i_2 i_3}, \\
\text{(b)} & \rho\pi'_{i_1 i_3 i_4} = \pm\, \vartheta_{i_2 i_3 i_4} & \cdot\ \Pi'/\vartheta_{i_2}\vartheta_{i_1 i_2 i_3}\vartheta_{i_1 i_2 i_4}, \\
\text{(c)} & \rho\pi'_{i_2 i_3 i_4} = \pm\, \vartheta_{i_1 i_3 i_4} & \cdot\ \Pi'/\vartheta_{i_1}\vartheta_{i_1 i_2 i_3}\vartheta_{i_1 i_2 i_4}, \\
\text{(d)} & \rho\pi'_{i_3 i_4 i} = \pm\, \vartheta_{i_1 \ldots i_5} & \cdot\ \Pi'/\vartheta_{i_1 i_2 i_3}\vartheta_{i_1 i_2 i_4}\vartheta_{i_1 i_2 i}\,, \\
& \multicolumn{2}{l}{\Pi' = \vartheta_{i_1}\vartheta_{i_2}\vartheta_{i_1 i_2 i_3} \cdots \vartheta_{i_1 i_2 i_{2p+2}}.}
\end{array}$$

From the following form of the equations (a)

$$\rho\pi'_{i_1 i_2 i_r} = \pm (1/\pi_{i_1 i_2 i_r}) \cdot \Pi'\Pi/\vartheta_{i_1}{}^2\vartheta_{i_2}{}^2 \qquad (r = 3, \cdots, 2p+2), \tag{a$'$}$$

we see that x' is the image of x in the involutorial regular transformation of order $2p-1$ whose $2p$ F-points are at $p_{i_3}, \cdots, p_{i_{2p+2}}$. If the factors in Π' with three indices are expressed in terms of the $\pi_{i_1 i_2 i_3}$, equations (a), (b), (c) become

$$\begin{array}{lll}
\text{(a}''\text{)} & \rho\pi'_{i_1 i_2 i_3} = \pm (\vartheta_{i_1}\vartheta_{i_2}/\Pi)^{2p-2} & \cdot\ \pi_{i_1 i_2 i_3}\pi_{i_1 i_2 i_4} \cdots \pi_{i_1 i_2 i_{2p+2}}/\pi_{i_1 i_2 i_3}, \\
\text{(b}''\text{)} & \rho\pi'_{i_1 i_3 i_4} = \pm (\vartheta_{i_1}\vartheta_{i_2}/\Pi)^{2p-2} & \cdot\ \pi_{i_2 i_3 i_4}\pi_{i_1 i_2 i_5} \cdots \pi_{i_1 i_2 i_{2p+2}}, \\
\text{(c}''\text{)} & \rho\pi'_{i_2 i_3 i_4} = \pm (\vartheta_{i_1}\vartheta_{i_2}/\Pi)^{2p-2} & \cdot\ \pi_{i_1 i_3 i_4}\pi_{i_1 i_2 i_5} \cdots \pi_{i_1 i_2 i_{2p+2}}.
\end{array}$$

Thus linear spreads on p_{i_2} become spreads of order $2p-1$ on p_{i_1}, and vice versa; whence the involution is the element $I_{i_1 i_2}$ of G_2^{n+2} $(n=2p-1)$. Since $I_{i_1 i_2}$ leaves w_p unaltered and on w_p is represented by $u' \equiv u + P_{i_1 i_2}$, then the symmetrical element $I = I_{i_1 \dots i_{2p+2}}$ of G_2^{n+2} leaves w_p unaltered and on w_p is represented by $u' \equiv u + P_{i_1 \dots i_{2p+2}} \equiv u$. Hence w_p is the locus of fixed points of I and coincides with W_p. The corresponding proof for the case p odd can be formulated by a similar procedure. Hence

(60) *The equations* (42), (43) *furnish the parametric equations of the generalized Weddle manifold W_p in S_{2p-1}. If only those equations* (42), (43) *which omit the index $2p+2$ in $\pi_{i_1 i_2 i_3}$ are retained, they furnish the parametric equations of W_p' in S_{2p-2}, the projection of W_p from p_{2p+2}. On W_p and W_p' the equation of the involution $I_{i_1 \dots 2l}$ is, in parametric form, $u' \equiv u + P_{i_1 \dots 2l}$.*

16. Mapping of W_p upon the generalized hyperelliptic Kummer manifold K_p. We revert to the formulae (47) for the transcendental equations of the P-loci of P^{2p-1}_{2p+2} and find that

$$(61) \qquad \begin{aligned} &p \text{ even}: \quad \pi_{i_1 \dots i_{2k+1}} \pi_{i_{2k+2} \dots i_{2p+2}} = \vartheta^2{}_{i_1 \dots i_{2k+1}} \cdot \Pi^{p-1}; \\ &p \text{ odd}: \quad \pi_{i_1 \dots i_{2k+1}} \pi_{i_{2k+2} \dots i_{2p+1}, 2p+2} = \vartheta^2{}_{i_1 \dots i_{2k+1}, 2p+2} \cdot \Pi^{p-1}. \end{aligned}$$

On the left we have the pairs (30) whose products are found in the linear system Σ of (31a). On the right we have, to within the factor of proportionality Π^{p-1}, the squares of the 2^{2p} odd and even functions of the first order. Of these only 2^p are linearly independent and an independent set may be taken to be coördinates of a point in the linear space S_{2^p-1}. For variable u this point runs over the manifold K_p, whence

(62) *The equations* (61) *represent a mapping of W_p upon K_p by means of the linear system Σ of spreads of order p with $(p-1)$-fold points at P^{2p-1}_{2p+2}. The section of W_p by either member of a pair of P-loci is mapped upon the section of K_p by one of its 2^{2p} singular spaces.*

We inquire as to the behavior of the singular points of K_p with reference to W_p. As is known, these occur on K_p when u is a half period. With respect to W_p we observe that the F-locus of the p-th kind of P^{2p-1}_{2p+2}, $\pi^{(p)}{}_{i_1 \dots i_p} = \pi^{(p)}{}_{i_{p+1} \dots i_{2p+2}}$, which is the linear space S_{p-1} on $p_{i_1}, \cdots, p_{i_p}$, is cut by the system Σ in a system of p-ic spreads with $(p-1)$-fold points at $p_{i_1}, \cdots, p_{i_p}$. But such a spread in S_{p-1} is unique. Hence the entire space S_{p-1} is mapped upon a single point of K_p which we proceed to identify with a singular point.

Let the hyperelliptic curve H_p^{p+2} satisfy the p further projective con-

ditions that $p+2$ of the branch points, $r_{i_{p+1}}, \cdots, r_{i_{2p+2}}$ are on a line L. Then the remaining p branch points are p flex-points at the p-fold point O, and the p-ad at O coincides with its superposed p-ad. Thus the parameters u of this p-ad must be a half period. When this situation arises, a number of the discriminant conditions are satisfied, and indeed, in many cases, are satisfied in more than one way.

Let the indices $j_1, j_2, \cdots$ be drawn from $i_{p+1}, \cdots, i_{2p+2}$. Of the two alternative forms (35) of the same discriminant condition we shall choose that which has more indices j than indices i $(i_1, \cdots, i_p)$. In

$$\delta(O^k j_1 \cdots j_{k+b+3} i_1 \cdots i_{k-b})^{k+1} = 0$$

this requires that $k+b+3 > k-b$. Also $1 \gtreqless k+b+3 \gtreqless p+2$ and $0 \gtreqless k-b \gtreqless p$. Hence

$$b > -3/2, \quad b \gtreqless k, \quad b \gtreqless p-k-1.$$

This permits of arranging the 2^{2p} distinct types of discriminant conditions in the order

$$\begin{array}{lll} & \text{(a)}: & b=-1;\ k=0,1,\cdots,p-1; \\ (63) & \text{(b)}: & b=b;\ k=-b, b+1,\cdots,p-1-b \\ & & \qquad [b=0,1,\cdots,(p-2)/2 \text{ or } (p-1)/2]. \end{array}$$

The line L must factor from the curve implied by the condition since

$$k+b+3 = (k+1)+(b+2) > k+1.$$

The remaining part must be a curve of order k with k-fold point at O and on $r_{i_1}, \cdots, r_{i_{k-b}}$. This part therefore consists of $k-b$ fixed lines on O and b variable lines on O. Hence if $b=-1$, the discriminant condition is not satisfied, whereas if $b>-1$ it is satisfied in $b+1$ linearly independent ways.

The point x on W_p which corresponds to this H_p^{p+2} with $r_{i_{p+1}}, \cdots, r_{i_{2p+2}}$ on L must (cf. (36)) lie on all the linear P-loci $\pi_{j_1 j_2 j_3}$ [for $b=0$, $k=0$ in (63b)], and therefore must be a point of the linear F-locus of the p-th kind $\pi^{(p)}{}_{i_1 \ldots i_p}$. To the discriminant condition $\delta(O^k j_1 \cdots j_{k+b+3} i_1 \cdots i_{k-b})^{k+1} = 0$ there correspond the conditions that x lies on $\pi_{j_1 \ldots {}_{k+b+3} i_1 \ldots i_{k-b}}$ and on $\pi_{j_{k+b+4} \ldots j_{p+2} i_{k-b+1} \ldots i_p}$. The first of these two P-loci consists (cf. (17)) of the ∞^k linear spaces S_{2p-k-1} which are $(2p-k)$-secant to N^{2p-1} at $p_{j_{k+b+4}}, \cdots, p_{j_{p+2}} p_{i_{k-b+1}}, \cdots, p_{i_p}$ and at k variable points of N^{2p-1}. If, of these k variable points, $k-b$ are taken at $p_{i_1}, \cdots, p_{i_{k-b}}$, there must be ∞^b of the spaces S_{2p-k-1} on $\pi^{(p)}{}_{i_1 \ldots i_p}$; and thus $\pi^{(p)}$ is a $(b+1)$-fold locus on π. The second of the two P-loci consists of the ∞^{p-k-2} linear spaces S_{p+k} which are $(p+k+1)$-secant to N^{2p-1} at $p_{j_1}, \cdots, p_{j_{k+b+3}}, p_{i_1}, \cdots, p_{i_{k-b}}$ and at $p-k-2$

variable points of N^{2p-1}. None of these spaces S_{p+k} can contain the further $p+b-k$ points p_i on $\pi^{(p)}$ whence the P-locus does not contain $\pi^{(p)}$. Hence

(64) *The F-locus of the p-th kind,* $\pi^{(p)}{}_{i_1\ldots i_p}=\pi^{(p)}{}_{i_{p+1},\ldots,i_{2p+2}}$, *is a* $(b+1)$-*fold locus of the P-locus of the first kind* $\pi_{j_1\ldots j_{k+b+3}i_1\ldots i_{k-b}}$ *(b and k chosen as in* (63)*), and is not on the complementary P-locus. Thus it is a* $(b+1)$-*fold locus on the corresponding product of the system* (61). *When x on* W_p *is also on* $\pi^{(p)}$, *the* $p+2$ *branch points,* $r_{j_1},\cdots,r_{j_{p+2}}$, *of* $H_p{}^{p+2}$ *are on a line.*

We have yet to identify this F-locus, $\pi^{(p)}{}_{i_1\ldots i_p}$, with a *particular* half period. The following theorem (cf. [8] p. 460) states the behavior of the hyperelliptic odd and even thetas with respect to the half periods, when these are expressed in the basis notation.

(65) *If the two sets of indices* $i_1\cdots i_{2l}$ *and* $j_1j_2\cdots$ *contain* $p+1\pm 2r$ *indices which do not appear in both sets, then if* $r>0$ *the theta function* $\vartheta_{j_1j_2}\cdots$ *vanishes, together with its derivatives up to the order* $r-2$ *inclusive, for the half period* $P_{i_1\ldots i_{2l}}$.

The non-vanishing for $r=0$ is normal, i. e., is common to all theta functions, as well as the vanishing of the functions for $r=1$; but the further evanescences are characteristic of the hyperelliptic functions.

When p is even we have identified in (42) the P-locus $\pi_{j_1\ldots j_{k+b+3}i_1\ldots i_{k-b}}$ with the theta function $\vartheta_{j_1\ldots j_{k+b+3}i_1\ldots i_{k-b}}$. By using the rule (65) we find that this function does not vanish for the half period $P_{i_1\ldots i_p}$ if $b=-1$ but does vanish for $b=0$, $b=1$. For values $b>1$ it vanishes together with its derivatives up to and including the order $b-1$. Hence $\pi^{(p)}{}_{i_1\ldots i_p}$ is to be identified with the half period $P_{i_1\ldots i_p}$. By applying the operations (27) and (60) we find that the half period $u=0$ is to be identified with the F-locus, $\pi^{(p)}$, the locus of the $\infty^{p/2}$ $S_{(p-2)/2}$'s on $p/2$ variable points of N^{2p-1}.

When p is odd, the P-locus $\pi_{j_1\ldots j_{k+b+3}i_1\ldots i_{k-b}}$ is identified with $\vartheta_{j_1\ldots j_{k+b+3}i_1\ldots i_{k-b},2p+2}$. If then the half period $P_{i_1\ldots i_p,2p+2}$ is identified with the F-locus, $\pi^{(p)}{}_{i_1\ldots i_p}$, the behavior of the P-loci and the theta functions with respect to respectively the F-locus and the half period coincides with that just described. This ascribes to the improper half period, $u=0$, the F-locus, $\pi^{(p)}{}_{2p+2}$, the locus of $\infty^{(p-1)/2}$ $S_{(p-1)/2}$'s on p_{2p+2} and $(p-1)/2$ variable points of N^{2p-1}. Hence

(66) *The* 2^{2p} *half periods determine on* W_p *the* 2^{2p} *F-loci of the p-th kind as follows:*

$$p\text{ even: }\quad \pi^{(p)}{}_{i_1\ldots i_{2k+2-p}}=\pi^{(p)}{}_{i_{2k+3-p}\ldots i_{2p+2}}\quad:\quad P_{i_1\ldots i_{2k+2-p}};$$
$$p\text{ odd: }\quad \pi^{(p)}{}_{i_1\ldots i_{2k+2-p}}=\pi^{(p)}{}_{i_{2k+3-p}\ldots i_{2p+2}}\quad:\quad P_{i_1\ldots i_{2k+2-p},2p+2}.$$

These F-loci are mapped as in (62) into the 2^{2p} singular points of K_p.

The mapping (62) furnishes a purely algebraic definition of the Kummer manifold K_p since W_p itself is the locus of fixed points of a Cremona involution defined by P^{2p-1}_{2p+2}. This algebraic definition suggests properties of K_p which may not readily appear from its transcendental definition. We examine in (83) one such situation for $p=3$.

§ 4

The Spread W_3' in S_4.

17. Theta relations in the hyperelliptic case, $p=3$. In the present case the functions are hyperelliptic when $\vartheta(0)=c=c_{12\ldots 8}=0$. The three-term relations are obtained from those four-term relations connecting the four products of four pairs of odd functions in a Steiner complex [1 p. 168, (3)]. With reference to the basis notation there are two types of Steiner complexes corresponding to the two types of half periods, P_{78} and P_{5678}. In the second type the four pairs can be selected in two distinct ways so that three types of four-term relations exist which connect the functions indicated:

$$\begin{array}{lllll}
\mathrm{A}: & \vartheta_{17}\vartheta_{18}, & \vartheta_{27}\vartheta_{28}, & \vartheta_{37}\vartheta_{38}, & \vartheta_{47}\vartheta_{48}; \\
\mathrm{B}: & \vartheta_{23}\vartheta_{14}, & \vartheta_{31}\vartheta_{24}, & \vartheta_{12}\vartheta_{34}, & \vartheta_{56}\vartheta_{78}; \\
\mathrm{C}: & \vartheta_{31}\vartheta_{24}, & \vartheta_{12}\vartheta_{34}, & \vartheta_{56}\vartheta_{78}, & \vartheta_{57}\vartheta_{68}.
\end{array}$$

The coefficients of the functions in these relations are, to within a $\pm$ sign:

$$\begin{array}{lllll}
\mathrm{A}': & c_{2347}c_{2348}, & c_{3147}c_{3148}, & c_{1247}c_{1248}, & c_{1237}c_{1238}; \\
\mathrm{B}': & c_{2378}c_{1478}, & c_{3178}c_{2478}, & c_{1278}c_{3478}, & cc_{1234}; \\
\mathrm{C}': & c_{1258}c_{3458}, & c_{3158}c_{2458}, & c_{5714}c_{6814}, & c_{5614}c_{7814}.
\end{array}$$

When $c=0$, Thomae's values (52) of the c_{ijkl} enable us to express these coefficients in terms of the roots of the fundamental octavic as follows:

$$\begin{array}{ll}
\mathrm{A}'': & [\{234\}\{156\}]^{1/2},\ [\{314\}\{256\}]^{1/2},\ [\{124\}\{356\}]^{1/2},\ [\{123\}\{456\}]^{1/2}; \\
\mathrm{B}'': & [(23)(14)]^{1/2},\ [(31)(24)]^{1/2},\ [(12)(34)]^{1/2},\ 0; \\
\mathrm{C}'': & [(12)(34)(58)(67)]^{1/2},\ [(31)(24)(58)(67)]^{1/2}, \\
& \qquad [(14)(23)(57)(68)]^{1/2},\ [(14)(23)(56)(78)]^{1/2}.
\end{array}$$

By adding a half period to u in these relations other relations of different form are obtained. We are concerned primarily with the three-term relations obtained from B, B″. The three terms are symmetric with respect to a division of the indices, 1234, 5678; and the addition of P_{1234} leaves each

term invariant. To secure typical results it is thus sufficient to add the half periods, P_{78}, P_{48}, P_{34}, P_{3478}. We secure the following five sets of three functions, each in a relation with coefficients B″:

B1:	$\vartheta_{23}\vartheta_{14}$,	$\vartheta_{31}\vartheta_{24}$,	$\vartheta_{12}\vartheta_{34}$;
B2:	$\vartheta_{2378}\vartheta_{1478}$,	$\vartheta_{3178}\vartheta_{2478}$,	$\vartheta_{1278}\vartheta_{3478}$;
B3:	$\vartheta_{2348}\vartheta_{18}$,	$\vartheta_{3148}\vartheta_{28}$,	$\vartheta_{1248}\vartheta_{38}$;
B4:	$\vartheta_{24}\vartheta_{13}$,	$\vartheta_{14}\vartheta_{23}$,	$\vartheta_{1234}\vartheta$;
B5:	$\vartheta_{2478}\vartheta_{3178}$,	$\vartheta_{1478}\vartheta_{2378}$,	$\vartheta_{56}\vartheta_{78}$.

Since, for the odd value of p, the index $2p+2=8$ is isolated in the application we are about to make, the relations just mentioned yield others as follows:

B1b:	$\vartheta_{23}\vartheta_{18}$;	$\vartheta_{31}\vartheta_{28}$,	$\vartheta_{12}\vartheta_{38}$	: (48);
B2b:	$\vartheta_{1256}\vartheta_{3456}$,	$\vartheta_{1257}\vartheta_{3457}$,	$\vartheta_{1267}\vartheta_{3467}$	: (17) (826354);
B3b:	$\vartheta_{1567}\vartheta_{14}$,	$\vartheta_{2567}\vartheta_{24}$,	$\vartheta_{3567}\vartheta_{34}$	: (48);
B2c:	$\vartheta_{2345}\vartheta_{15}$,	$\vartheta_{1345}\vartheta_{25}$,	$\vartheta_{1245}\vartheta_{35}$	: (58);
B3d:	$\vartheta_{1234}\vartheta_{18}$,	$\vartheta_{2567}\vartheta_{12}$,	$\vartheta_{3567}\vartheta_{13}$	: (81);
B4b:	$\vartheta_{56}\vartheta_{78}$,	$\vartheta_{57}\vartheta_{68}$,	$\vartheta_{1234}\vartheta$	: (48) (27) (16) (35);
B5b:	$\vartheta_{1256}\vartheta_{3456}$,	$\vartheta_{1257}\vartheta_{3457}$,	$\vartheta_{12}\vartheta_{34}$	: (71825463).

The substitution (in cycle form) on the right of each is that which converts Bi into the given form; and which therefore must be applied also to the coefficients B'' to obtain the coefficients of the given terms.

The three-term relations have been derived afresh to show their place in the relations among the non-hyperelliptic functions. Thy might have been obtained directly from (44b).

18. The theta relations of 17 as projective relations in S_4. The P-loci of the set $P_7{}^4$ in S_4, π_{i_1}, $\pi_{i_1i_2i_3}$, $\pi_{i_1\ldots i}$, and $\pi_{i_1\ldots i_7}$ $(i_j=1,\cdots,7)$ are defined in (47) in terms of the thetas:

$$\pi_{i_1\ldots i_{2k+1}}=\vartheta_{8i_1\ldots i_{2k+1}}\cdot\Pi^k/\vartheta_{8i_1}\ldots\vartheta_{8i_{2k+1}}\qquad(k=0,\cdots,3),\qquad \Pi=\vartheta_{81}\vartheta_{82}\cdots\vartheta_{87}. \tag{67a}$$

Three-term relations among products of these P-loci are then derived from the theta relations. We have noticed already in (59), (57) an example of a projective relation in S_4 which is obtained from a theta relation. We now extend the definition of $G_{i_1i_2i_3}$ in (59) to cover all the P-loci of $P_7{}^4$ by setting

$$G_{i_1\ldots i_{2k+1}}=\pi_{i_1\ldots i_{2k+1}}\cdot\{i_1\cdots i_{2k+1}\}^{1/2}. \tag{67b}$$

The factors, $\{i_1\cdots i_{2k+1}\}$, are defined in (49) in terms of the differences

(ij) of the roots t_i, t_j $(i, j = 1, \cdots, 7)$ of the fundamental octavic. These differences are proportional to the quinary determinants formed from the coördinates of the points of $P_7{}^4$ as in (50); specifically

$$(67c) \qquad (i_1 \cdots i_5) = \lambda \cdot \epsilon_{i_1 \ldots i_5 i_6 i_7} (i_6 i_7).$$

In order that a relation involving these differences may be interpreted as a projective relation, it is necessary that a certain number (the same for each term) of the differences be interpreted as quinary determinants so that the terms may be homogeneous in each point of $P_7{}^4$; and that the products of the remaining differences in each term be homogeneous in each of the roots $t_1, \cdots, t_8$ so that the ratios of the products are invariants, rational or irrational, of the octavic. An example is the transition from (57) to the preceding purely projective relation.

We prove that the functions G satisfy the same system of relations as the projective expressions in (67d) and therefore may be identified with them.

$$(67d) \quad \begin{aligned} G_{123} &= \pm (4567x) \\ G_{1 \ldots 5} &= \pm [(23)(14)(2467x)(1367x) - (13)(24)(1467x)(2367x)], \\ G_{1 \ldots 7} &= \pm [(23)(14) G_{13567} G_{24567} - (13)(24) G_{23567} G_{14567}]. \end{aligned}$$

Thus the respective degrees of G_{123}, $G_{1 \ldots 5}$, and $G_{1 \ldots 7}$ in the coördinates of $P_7{}^4$ are

$$(0\ 0\ 0\ \ 1\ 1\ 1\ 1), \qquad (2\ 2\ 2\ 2\ 2\ 4\ 4), \qquad (6\ 6\ 6\ 6\ 6\ 6\ 6).$$

Multiply B3 in **17** by $\Pi/\vartheta_{18} \cdots \vartheta_{48}$ to introduce the functions π in (67a); with coefficients B'', then insert the values G from (67b); and multiply by $\{4; 123\}^{1/2}$ to remove the denominators. The resulting relation is

$$(14) G_{234} \pm (24) G_{314} \pm (34) G_{124} \equiv 0.$$

With the differences replaced by quinary determinants this is the projective relation satisfied by S_3's on p_5, p_6, p_7 to p_1, p_2, p_3 respectively, and justifies the first formula of (67d).

Again we multiply in B5 by $\Pi^2/\vartheta_{18} \cdots \vartheta_{48} \vartheta^2{}_{78}$ to introduce the functions π, and after inserting the coefficients B'' and the functions G, multiply by $[\{12347\}/(12)(34)]^{1/2}$ to obtain

$$(23)(14) G_{247} G_{317} \pm (24)(13) G_{147} G_{237} \pm G_{12347} \equiv 0.$$

On interchanging 7 and 5, this is the second formula (67d). The determination of the relative sign of the first two terms given in (67d) is based on the fact that the third term, $G_{12347} = \rho \pi_{12347}$, must represent a quadric with nodes at p_5, p_6 and on $p_1, \cdots, p_4, p_7$.

Finally on multiplying B4 by $\Pi^3/\vartheta_{58}\vartheta_{68}\vartheta^2_{78}$, and by

$$[\{567\}\{1234567\}/(12)(34)]^{1/2},$$

we convert it into the projective relation,

$$(23)(14)G_{13567}G_{24567}-(13)(24)G_{23567}G_{14567}=\pm G_{567}G_{1\ldots7},$$

which is the third formula of (67d). We verify in a moment that

$$(67e)\quad G_{1234567}=\pm[(14)(46)(63)(32)(25)(51)(1342x)(1365x)(4265x)\\(14)(42)(23)(36)(65)(51)(1346x)(1325x)(4625x)\\(14)(46)(65)(52)(23)(31)(1542x)(1563x)(4263x)\\(13)(36)(64)(42)(25)(51)(1432x)(1465x)(3265x)\\(13)(36)(65)(52)(24)(41)(1532x)(1564x)(3264x)\\(13)(32)(24)(46)(65)(51)(1436x)(1425x)(3625x)].$$

We complete the examination of the relations, $B1,\cdots,B5$, by multiplying $B1$, $B2$ respectively by

$$[\{567;\ 1234\}\{567\}^2]^{1/2}\,\Pi^3/\vartheta_{58}\vartheta_{68}\vartheta_{78},$$
$$[\{7;\ 1234\}]^{1/2}\,\Pi^2/\vartheta_{18}\cdots\vartheta_{48}\vartheta^2_{78}$$

to reduce them to

$$(67f)\quad \begin{array}{l}G_{14567}G_{23567}\pm G_{24567}G_{31567}\pm G_{34567}G_{12567}\equiv 0,\\ G_{237}G_{147}\pm G_{317}G_{247}\pm G_{127}G_{347}\equiv 0.\end{array}$$

The second of these two is the identity which arises when the determinants in the ordinary binary identity, $(14)(23)+(24)(31)+(34)(12)=0$, are expanded into the quinary domain by the insertion of $56x$ in each determinant according to the Clebsch principle of transference. The first is the form which the second takes when the Cremona involution I_{56} is applied to it. A more direct proof is not immediate so that this identity is an example of the usefulness of the thetas in a projective study of $P_7{}^4$ and related loci.

A similar example is furnished by B3c which yields

$$(67g)\qquad (14)G_{167}G_{23467}\pm(24)G_{267}G_{31467}\pm(34)G_{367}G_{12467}\equiv 0.$$

The Cremona involution I_{46} applied to the second relation (67f) will produce an identity of this character.

The transcription of the relations B1b, $\cdots$, B5b and B3d is not so direct and will be deferred until more information concerning W_3' is gained (cf. **19**). We close this section with an examination of $G_{1\ldots7}$ in (67e), the interesting locus of the bisecants of the curve N^4 on $P_7{}^4$ (cf. **4**, table). If in (67d) we take the quadrics in the form,

$$G_{13567} = (2413x)(2456x)(15)(36) - (2415x)(2436x)(13)(56),$$
$$G_{24567} = (1324x)(1356x)(25)(46) - (1325x)(1346x)(24)(56),$$
$$G_{23567} = (1423x)(1456x)(25)(36) - (1425x)(1436x)(23)(56),$$
$$G_{14567} = (2314x)(2356x)(15)(46) - (2315x)(2346x)(14)(56),$$

then the combination, $(23)(14)G_{13567}G_{24567} - (13)(24)G_{23567}G_{14567}$ contains the factor, $G_{567} = (1234x)$. Indeed the only two terms of the combination, which do not contain this factor explicitly, cancel each other. The complementary factor is the bracket in (67e). Each of the terms in $G_{1\ldots 7}$ involves one of the cyclic arrangements:

146325, 142365, 146523, 136425, 136524, 132465.

For each cycle the pairs of diagonals of the corresponding hexagon determine the quinary determinants. The six cycles may be described as the *six ways in which* 126 *can be inserted between* 345 *or vice versa.* Thus there must be ten formally different equations (67e) of $G_{1\ldots 7}$; or $10.7 = 70$ if also the exceptional position of p_7 be considered.

It is evident from the given form that the cubic spread $G_{1\ldots 7}$ has nodes at $p_1, \cdots, p_6$. That it has a double point at p_7 is also easy to see. For, the second polar of p_7 is obtained by setting $x = p_7$ in each pair of terms in each of the six products of three quinary determinants. Then any determinant, e.g. $(1234x)$, appears twice with coefficients which cancel when (13567) is replaced by (24), etc. On account of the symmetry it is sufficient to verify this for $(1234x)$ alone. Since p_7 does not appear explicitly in the equation the locus must have a node at every point $t = t_7$ on N^4, and therefore must be the locus of bisecants of N^4.

19. Geometric properties of W_3'. We have seen [cf. (66)] that the F-loci of the third kind of P_8^5 in S_5 are contained on W_3. When W_3 is projected from p_8 upon W_3' in S_4 these F-loci of the third kind of P_8^5 are projected into the F-loci of the second kind of P_7^4 (cf. **4**, table). Hence W_3' must contain these F-loci of the second kind of P_7^4 for the values of the half periods as follows:

$$\begin{array}{lcl} u = 0 & : & \pi^{(2)} = N^4; \\ u = P_{i_1 i_2} & : & \pi^{(2)}{}_{i_1 i_2} = S_1(i_1 i_2)^1; \\ u = P_{i_1 i_2 i_3 i_4} & : & \pi^{(2)}{}_{i_1 i_2 i_3 i_4} = S_2(i_5 i_6 i_7)^1; \\ u = P_{i_7 8} & : & \pi^{(2)}{}_{i_1 \ldots i_7} = C^3{}_{i_7} = M_2(i_1 \cdots i_6 i_7^3)^3. \end{array}$$

Since W_3' is invariant under the G_{64} determined by P_7^4 it must be a spread of order $5k$ with $3k$-fold points at P_7^4. Moreover since the same pro-

jective set P_7^4 can be obtained by the projection of a P_8^5 in S_5 which consists of P_7^5 and any point t on the N^5 determined by P_7^5 and p_8, the equation of W_3' in S_4 can involve $t = t_8$ only in the coefficients. Hence W_3' will be a member of a system (∞^1) determined by P_7^4. We shall find that this system is a pencil.

The projective relations thus far obtained from the theta relations have been identities among the P-loci which were valid throughout the space. These could lead to no definition of W_3'. We consider then the further set B1b, $\cdots$, B5b. In order to handle the functions $\vartheta^2_{18}, \cdots, \vartheta^2_{78}$, which do not disappear as in the earlier cases, we define two new sets of seven functions, $H_1, \cdots, H_7$; $H_1', \cdots, H_7'$, as follows:

$$
\begin{aligned}
H_1 &= \vartheta^2_{18} \cdot w\Pi^2 \cdot [\{234567\}(18)]^{1/2}, \\
H_1' &= [1/\vartheta^2_{18}] \cdot w\Pi^4 \cdot [\{1234567\}\{1;\ 234567\}(18)]^{1/2},
\end{aligned} \tag{68a}
$$

where w is a function to be determined later. Hence

$$
\begin{gathered}
H_1H_1' = (18) \cdot \Pi^6 w^2 \cdot \{1234567\}, \cdots, H_7H_7' = (78) \cdot \Pi^6 w^2 \cdot \{1234567\}, \\
(23)H_1H_1' + (31)H_2H_2' + (12)H_3H_3' \equiv 0.
\end{gathered} \tag{68b}
$$

In terms of the functions H, H' the later theta relations transcribe as follows:

$$
\begin{aligned}
&\text{B1b:} \quad G_{14567}H_1 \pm G_{24567}H_2 \pm G_{34567}H_3 \equiv 0; \\
&\text{B2b:} \quad G_{347}G_{127}H_7 \pm G_{346}G_{126}H_6 \pm G_{345}G_{125}H_5 \equiv 0; \\
&\text{B3b:} \quad G_{234}G_{23567}H_1' \pm G_{314}G_{31567}H_2' \pm G_{124}G_{12567}H_3' \equiv 0; \\
&\text{B4b:} \quad (57)G_{12347}H_6' \pm (56)G_{12346}H_7' \pm G_{567}G_{1\ldots7}H_5 \equiv 0; \\
&\text{B5b:} \quad (57)G_{127}G_{347}H_6' \pm (56)G_{126}G_{346}H_7' \pm G_{12567}G_{34567}H_5 \equiv 0.
\end{aligned} \tag{68c}
$$

If B4b and B5b are multiplied by H_6H_7, they reduce by using (68b) to

$$
\begin{aligned}
&\text{B4b}': \quad (57)(68)G_{12347}H_7 \pm (56)(78)G_{12346}H_6 \\
&\qquad\qquad \pm G_{567}G_{1\ldots7}H_5H_6H_7/\Pi^6w^2\{1\cdots7\} \equiv 0, \\
&\text{B5b}': \quad (57)(68)G_{127}G_{347}H_7 \pm (56)(78)G_{126}G_{346}H_6 \\
&\qquad\qquad \pm G_{12567}G_{34567}H_5H_6H_7/\Pi^6w^2\{1\cdots7\} \equiv 0.
\end{aligned} \tag{68d}
$$

We write the identity B2b for the three combinations, 12, 34; 13, 42; and 14, 23, using the projective form except in the unknown H_5, H_6, H_7. Thus

$$
\begin{aligned}
&\pm(1256x)(3456x)H_7 \pm (1257x)(3457x)H_6 + (1267x)(3467x)H_5 \equiv 0, \\
&\pm(1356x)(4256x)H_7 \pm (1357x)(4257x)H_6 + (1367x)(4267x)H_5 \equiv 0, \\
&\pm(1456x)(2356x)H_7 \pm (1457x)(2357x)H_6 + (1467x)(2367x)H_5 \equiv 0.
\end{aligned}
$$

If these be added the terms in H_5 vanish due to the second identity (67f). Since H_6 and H_7 with significant factors $\vartheta^2{}_{68}$, $\vartheta^2{}_{78}$ could not remain, the terms containing them must also vanish due to the same identity, whence the ambiguous signs are alike in each column. Any two of the three identities then serve to determine the ratios of H_5, H_6, H_7. We shall set

$$(68e)\qquad H_7 = \pm \begin{vmatrix} (1257x)(3457x) & (1267x)(3467x) \\ (1357x)(4257x) & (1367x)(4267x) \end{vmatrix} = \begin{vmatrix} G_{346}G_{126} & G_{345}G_{125} \\ G_{426}G_{136} & G_{425}G_{135} \end{vmatrix}.$$

This expression for H_7 determines the function w in (68a). For, if the quinary determinants are replaced from (67a, b, d), H_7 takes the form

$$(68f)\qquad H_7 = \pm\, \Pi^2 \cdot \vartheta^2{}_{78} \cdot [\{123456\}(78)]^{1/2} \cdot \frac{\vartheta_{3468}\vartheta_{1268}\vartheta_{2458}\vartheta_{1358} \pm \vartheta_{3458}\vartheta_{1258}\vartheta_{2468}\vartheta_{1368}}{[(14)(23)(56)(78)]^{1/2}}$$

On comparing this with (68a) we find that

$$(68g)\qquad w = (\vartheta_{3468}\vartheta_{1268}\vartheta_{2458}\vartheta_{1358} \pm \vartheta_{3458}\vartheta_{1258}\vartheta_{2468}\vartheta_{1368})/[(14)(23)(56)(78)]^{1/2}.$$

We recognize at once from (68e) that H_7 is the locus of points x such that $p_1, \cdots, p_6$ are projected from the line p_7x into six points on a conic; or, it is the locus of points x which are projected from p_7 into the points of a Weddle quartic surface with nodes at the projections of $p_1, \cdots, p_6$. It is thus a Weddle quartic cone with 4-fold point at p_7 and double lines from p_7 to $p_1, \cdots, p_6$. It contains the remaining F-loci of the second kind and dimension one simply [i. e., N^4 and the 15 lines $S_1(i_1i_2)^1$ $(i_1, i_2 = 1, \cdots, 6)$]; and also contains F-loci of the second kind and dimension two; namely $C_7{}^3$, and the 15 planes $S_2(7i_1i_2)^1$.

The identity B1b brings out the significance of the factor w in the definition of the Weddle cones H. The two sextic spreads $G_{14567}H_1$ and $G_{24567}H_2$ lie in a pencil with a basic 2-way of order 36. The quadrics G_{14567} and G_{24567} meet in $C_3{}^3$ and $S_2(123)^1$; G_{14567} and H_2 meet in $C_2{}^3$ and $S_2(231)^1$, $S_2(234)^1, \cdots, S_2(237)^1$; G_{24567} and H_1 meet in $C_1{}^3$ and $S_2(132)^1$, $S_2(134)^1$, $\cdots, S_2(137)^1$; and H_1 and H_2 meet in $S_2(123)^1$, $S_2(124)^1, \cdots, S_2(127)^1$ and a $V_2{}^{11}$. Since the complete intersection of H_1 and H_2 is a 16-ic with 4-fold points at p_1, p_2 and 4-fold points at $p_3, \cdots, p_7$, $V_2{}^{11}$ must have triple points at $P_7{}^4$. Since $G_{34567}H_3$ is a member of this pencil, $V_2{}^{11}$ must also lie on H_3, and thus must lie on each of the seven Weddle cones. Also H_1, H_2 contain N^4 and the line $S_1(34)^1$. These not being on the 5 planes common to H_1, H_2 must be on $V_2{}^{11}$. On account of the symmetry of $V_2{}^{11}$ with respect to the seven cones we may state that

(69) *The seven Weddle quartic cones H_i with respectively vertices at p_i*

and double lines p_ip_j meet in a surface V_2^{11} with triple points at P_7^4 and containing N^4 and the 21 lines $S_1(ij)^1$. The surface V_2^{11} is invariant under the Cremona G_{64} defined by P_7^4. The equation of this surface on W_3' is $w=0$ (cf. 68g).

For, the Weddle cones, H_6 and H_7, are each invariant under the involutions generated by $I_{i_1i_2}$ $(i_1, i_2 = 1, \cdots, 5)$; whence V_2^{11}, their common part except for F-loci, is also invariant under these involutions. That V_2^{11} is on W_3' is a consequence of the equation of W_3' we are about to derive.

If in (68d) the third term of the two identities is eliminated, we find that

$$(70)\qquad \begin{aligned} W_3' = (57)(68)\cdot[G_{12567}G_{34567}G_{12347} - G_{567}G_{127}G_{347}G_{1\ldots 7}]\,H_7 \\ \pm(56)(78)\cdot[G_{12567}G_{34567}G_{12346} - G_{567}G_{126}G_{346}G_{1\ldots 7}]\,H_6 = 0. \end{aligned}$$

This relation (70) is an identity in u but it is not a projective identity. For, the S_3, $G_{567}=0$, is a factor of the two of the terms but not a factor of the other two for variable t_8. We observe that each of the brackets is of degree 10 in each point p of P_7^4 as well as of degree 10 in x. Hence the factors (57)(68), (56)(78) are to be interpreted in the binary domain as terms of a double ratio. Thus W_3' appears, as observed earlier, as a member of a pencil (for variable t_8) defined by P_7^4.

The alternative forms of the function w of (68g) are 105 in number. These arise from the 15 choices of the form of H_7 in (68e) which depend upon the division of the six indices into pairs (14)(23)(56); and from the 7 choices of the index 7.

The relations B3b play the same role in the determination of the functions H' as the relations B2b in the determination of the H's. Let B3b be written three times for the cyclic permutation of 4, 5, 6. The coexistence of the resulting three identities is a consequence of the identities (67g) with multipliers (47), (57), (67). Any two of the identities serve to determine the ratios of H_1', H_2', H_3'. If we select the first two, an extraneous constant factor (67) appears, so that we set

$$(71a)\qquad (67)\cdot H_1' = \pm\begin{vmatrix} G_{134}G_{13567} & G_{124}G_{12567} \\ G_{135}G_{13467} & G_{125}G_{12467} \end{vmatrix}.$$

That this projective form of H_1' coincides with the definition of H_1' in (68a) appears when the spreads G are expressed in terms of the functions by means of (67a, b). The expression is

$$(71b)\qquad \begin{aligned} (67)\cdot H_1' = \pm(67)\cdot[\Pi^4/\vartheta^2_{18}]\cdot[\{1\cdots 7\}\{1;\, 2\cdots 7\}(18)]^{1/2} \\ \cdot(\vartheta_{24}\vartheta_{35}\vartheta_{1348}\vartheta_{1258} \pm \vartheta_{34}\vartheta_{25}\vartheta_{1248}\vartheta_{1358})/[(18)(23)(45)(67)]^{1/2}. \end{aligned}$$

This value of H_1' coincides with that in (68a) except for the form of the function w. By equating two alternative forms of w in (68g),

$$
\begin{aligned}
&(\vartheta_{3568}\vartheta_{1258}\vartheta_{1348}\vartheta_{2468} \pm \vartheta_{1358}\vartheta_{2658}\vartheta_{3468}\vartheta_{1248})/[(16)(23)(45)(78)]^{1/2} \\
&= \pm(\vartheta_{3578}\vartheta_{2478}\vartheta_{3468}\vartheta_{2568} \pm \vartheta_{2578}\vartheta_{3478}\vartheta_{2468}\vartheta_{3568})/[(18)(23)(45)(67)]^{1/2},
\end{aligned}
$$

we have an identity which, when u is replaced $u+P_{16}$, expresses the function in (71b) as one of the alternative forms of w. Thus (71a) is a projective definition of H_1'. According to it the degrees of H_1' in the points $p_1, \cdots, p_7$ are (3, 6, 6, 6, 6, 6, 6).

Other forms of H_7' occur; e. g., the coefficient of H_7 in (70). We find, again by using (67a, b), that this coefficient is

$$
\begin{aligned}
&(12)(34)(56)\cdot[\Pi^4/\vartheta^2{}_{78}]\cdot[\{1\cdots 7\}\{7;1\cdots 6\}(78)]^{1/2}. \\
&\qquad\cdot(\vartheta_{34}\vartheta_{12}\vartheta_{56}\vartheta_{78} - \vartheta_{5678}\vartheta_{1278}\vartheta_{3478}\vartheta)/[(12)(34)(56)(78)]^{1/2}.
\end{aligned}
$$

By the same argument this is $(12)(34)(56)\cdot H_7'$ whence

$$
(71c) \qquad (12)(34)(56)\cdot H_7' = \begin{vmatrix} G_{12347} & G_{567}G_{1\ldots 7} \\ G_{127}G_{347} & G_{12567}G_{34567} \end{vmatrix}.
$$

On replacing the coefficients of H_7 and H_6 in (70) by these values we find that

$$
(72) \quad W_3' = (12)(34)\cdot[(57)(68)\cdot(56)H_7H_7' \pm (56)(78)\cdot(57)H_6H_6'].
$$

We observe that both terms, $(56)H_7H_7'$ and $(57)H_6H_6'$, have degrees (9, 9, 9, 9, 8, 8, 8); and the factor (12)(34) raises all of these degrees to 10.

The spread, $H_7'=0$, is a sextic spread with four-fold points at $p_1, \cdots, p_6$ and double point at p_7, which is on V_2^{11}. According to its transcendental definition (68a), these properties are sufficient to define it. The Weddle quartic cone H_7 is invariant under the G_{32} whose involutions $I_{i_1\ldots i_{2k}}$ do not contain the index 7. It has therefore but one conjugate under G_{64}. The image of H_7 under I_{17} is a sextic spread with double point at p_7, 4-fold points at $p_1, \cdots, p_6$, and on V_2^{11} which is invariant under I_{17} [cf. (69)]. This image must coincide with H_7'. Hence each of the seven products H_iH_i' is invariant under G_{64}. According to (68b) these seven products are in a pencil. We observe that when, in (72), $t_8 = t = t_7$, W_3' reduces to H_7H_7'. Hence

(73) *Each of the seven Weddle cones H_i is invariant under a G_{32} in G_{64}, and is conjugate under G_{64} to a sextic spread H_i' with node at p_i and 4-fold points at the other 6 points p. The seven products H_iH_i' are in a pencil with parameters $t=t_i$. The spread W_3' is in this pencil with parameter*

$t = t_8$. *The pencil is that of order* 10, *which contains* V_2^{11} *as a double locus and the F-loci of the second kind of* P_7^4 *as a simple locus.*

The base of the pencil, a surface of order 100, consists of V_2^{11} counted four times, of the seven cubic cones C_i^3, and of the 35 planes $S_2(ijk)^1$. We say that the pencil contains the one-dimensional F-loci as *simple* loci if it contains them to a multiplicity one greater than they must occur on spreads of order 10 with 6-fold points at P_7^4. Thus $S_1(12)^1$ must be a double line of spreads of such a pencil. It is however a triple line on the pencil (70), and a like situation with respect to N^4 and the 21 lines $S_1(ij)$ must exist.

We give finally a determination of sign for the spread W_3'.

(74) *If the signs of* H_1, H_2, H_3 *and* H_1', H_2', H_3' *are so chosen that the identity* (68b) *reads:*

$$(23)H_1H_1' + (31)H_2H_2' + (12)H_3H_3' \equiv 0,$$

then the equation of W_3' *can be written as*

$$\left\| \begin{matrix} (23)H_1H_1' & (31)H_2H_2' & (12)H_3H_3' \\ (23)(18) & (31)(28) & (12)(38) \end{matrix} \right\| = 0.$$

For, if $x_0 + x_1 + x_2 = 0$ and $y_0 + y_1 + y_2 = 0$, then $x_1y_2 - x_2y_1 = x_2y_0 - x_0y_2 = x_0y_1 - x_1y_0$. Only with this disposition of the $\pm$ signs could the three forms of W_3', as given in (72), coexist.

20. The double surface V_2^{11} of W_3'. Let L be a generic line of the Weddle quartic cone H_7 with vertex at p_7. This is a fixed line of the involution I_{123456} whose parametric equation on W_3' and V_2^{11} is $u' \equiv u + P_{78}$. The line L meets H_6 in two points outside p_7 which are points on V_2^{11} and double points on W_3', whence L meets V_2^{11} and W_3' only at these two points outside p_7. These points must be interchanged by $I_{1\ldots 6}$, since $u' \equiv u + P_{78}$ has only 32 fixed points on W_3' which are determined by the quarter periods, $\pm P_{78}/2$. Since each point of V_2^{11} is on one line L, there follows that:

(75) *On* V_2^{11}, *invariant under* G_{64}, *the elements of* G_{64} *are generated by combining the involutions* I_{i8} $(i = 1, \cdots, 7)$ *which respectively are the projections of* V_2^{11} *into itself from* p_i. *Under successive projections from the triple points a point* u *of* V_2^{11} *gives rise to a closed set of* 64 *points conjugate under* G_{64}. *The surface* V_2^{11} *is projected from any one of its triple points, say* p_7, *into a doubly covered Weddle quartic surface with fundamental sextic* $t_1, \cdots, t_6$; *and the conjugate sets of* 64 *points on* V_2^{11} *are projected into sets of* 32 *points on the Weddle of the type described by Baker* [11] (*cf. also Hudson,*[12] p. 169).

This V_2^{11} is a particular case of a surface which occurs in connection with every set P^n_{n+3} $(n \geqq 3)$ in S_n. The rational curve of order $n+1$ in S_n, R^{n+1}, with a node can be put into the unique canonical form,

$$x_0 = (t_0^{n+1} + t_1^{n+1})/2, \quad x_i = t_0^{n+1-i} t_1^{i} \qquad (i = 1, \cdots, n).$$

It has therefore $n(n+2)$ constants. On P^n_{n+3} there are ∞^3 R^{n+1}'s with a node, since $n(n+2) - (n+3)(n-1) = 3$. If x is the node of such an R^{n+1}, the R^{n+1} is projected from x into an N'^{n-1} on the projected set P'_{n+3} in S_{n-1}. The cone of projection, K^{n-1}, has an $(n-1)$-fold point at x. Since N'^{n-1} is on $(n-1)(n-2)/2$ independent quadrics, K^{n-1} is contained on a system Σ' $(\infty^{-1+(n-1)(n-2)/2})$ of quadrics in S_n with a node at x. The system Σ $(\infty^{(n^2+n-8)/2})$ of quadrics in S_n on P^n_{n+3} and x is made up of Σ', and a system Σ'' (∞^{2n-5}) which does not contain K^{n-1}. The system Σ'' cuts K^{n-1} in a linear system of curves L'', of dimension $2n-5$, order $2n-2$, and with an $(n-1)$-fold point at x. The system L'' contains R^{n+1}'s with node at x together with $n-3$ arbitrary lines of K^{n-1}. It is two conditions on Σ'' to contain such a line whence there are ∞^1 R^{n+1}'s on K^{n-1}, which have a node at x and pass through P^n_{n+3}. Thus the locus of the node x of the ∞^3 nodal R^{n+1}'s on P^n_{n+3} is a surface. Hence

(76) *Given P^n_{n+3} there exists a surface $V_2^{(n)}$, the locus of a point x:* (a) *from which P^n_{n+3} projects into a set P'^n_{n+3} in S_{n-1} on a norm-curve N^{n-1};* (b) *which is the $(n-1)$-fold point of a cone K^{n-1} of order $n-1$ on P^n_{n+3};* and (c) *which is the node of ∞^1 of the rational curves R^{n+1} on P^n_{n+3}. This surface $V_2^{(n)}$ is the complete intersection of the spreads $H_{i_7 \ldots i_{n+3}}$, the locus of S_{n-3}'s on $p_{i_7}, \cdots, p_{i_{n+3}}$ to points x' of the Weddle surface determined in S_3' by the projections of $p_{i_1}, \cdots, p_{i_6}$ from the S_{n-4} on $p_{i_7}, \cdots, p_{i_{n+3}}$. The surface $V_2^{(n)}$ contains all the lines $p_i p_j$, and the curve N^n on P^n_{n+3}. It has $(n-1)$-fold points at P^n_{n+3} such that the tangent cone at p_{i_1} is the cone K^{n-1} from p_{i_1} to the N'^{n-1} on the projections of $p_{i_2}, \cdots, p_{i_{n+3}}$.*

For $n=3$ this theorem states the more fundamental properties of the Weddle surface; for $n=4$ it gives additional properties of the surface $V_2^{11} = V_2^{(4)}$.

21. Mapping of W_3' upon the Kummer manifold K_3 in S_7. We derive from the theta relations one such mapping without attempting to draw any geometric conclusions. The mapping is given by the equations:

$$
(77)\quad
\begin{aligned}
&\text{(a)}\ \{1\cdots 7\}[\{1\cdots 8\}]^{1/2}\cdot\vartheta^2\Pi^6 w\cdot\vartheta^2\\
&\quad=[\pm(17)\{8;\,234\}\cdot\{1;\,56\}H_1\pm\cdots\\
&\qquad\pm(47)\{8;\,123\}\cdot\{4;\,56\}H_4]\,G^2{}_{1\ldots 7};\\
&\text{(b)}\ \{1\cdots 7\}[\{123568\}(47)]^{1/2}\cdot\vartheta^2\Pi^6 w\cdot\vartheta^2{}_{47}\\
&\quad=[\pm(16)\{8;\,23\}\cdot(15)G_{147}H_1\pm\cdots\\
&\qquad\pm(36)\{8;\,12\}\cdot(35)G_{347}H_3]G_{12356}G_{1\ldots 7};\\
&\text{(c)}\ \{1\cdots 7\}[\{5678\}\{1234\}]^{1/2}\cdot\vartheta^2\Pi^6 w\cdot\vartheta^2{}_{5678}\\
&\quad=[\pm(57)(68)\cdot H_7G_{12347}\pm(56)(78)\cdot H_6G_{12346}]G_{567}G_{1\ldots 7};\\
&\text{(d)}\ \{1\cdots 7\}[\{2\cdots 7\}(18)]^{1/2}\cdot\vartheta^2\Pi^6 w\cdot\vartheta^2{}_{18}=H_1G^2{}_{1\ldots 7}.
\end{aligned}
$$

In these equations we have on the left the four types of theta squares, ϑ^2, $\vartheta^2{}_{47}$, $\vartheta^2{}_{5678}$, $\vartheta^2{}_{18}$ each multiplied by the factor $\vartheta^2\Pi^6 w$, and by a modular factor which varies somewhat from type to type. On the right the brackets properly interpreted represent projective loci. Thus in (a) the terms $\{1;\,56\}H_1=(15)(16)H_1,\cdots,\{4;\,56\}H_4$ are all of degrees (4, 4, 4, 4, 3, 3, 3) in $p_1,\cdots,p_7$ respectively when (15) is replaced by (23467), etc., and of degree 4 in x; whereas the coefficients $(17)\{8;234\},\cdots,(47)\{8;123\}$ have ratios which are double ratios of the fundamental octavic.

The equation (d) is a consequence of the definitions (67a, b), (68a). The equation (c) is obtained from the identity B4b′ in (68d). The equation (b) is a transcription of one of the identities A in **17** which reads:

$$
[\{123\}\{856\}]^{1/2}\,\vartheta\vartheta_{47}\pm[\{238\}\{156\}]^{1/2}\vartheta_{2356}\vartheta_{18}\pm[\{138\}\{256\}]^{1/2}\vartheta_{3156}\vartheta_{28}\pm[\{128\}\{356\}]^{1/2}\vartheta_{1256}\vartheta_{38}\equiv 0.
$$

The equation (a) is a transcription of a five-term identity:

$$
c^2{}_{1234}\vartheta^2\pm c^2{}_{2348}\vartheta^2{}_{18}\pm c^2{}_{1348}\vartheta^2{}_{28}\pm c^2{}_{1248}\vartheta^2{}_{38}\pm c^2{}_{1238}\vartheta^2{}_{48}\equiv 0.
$$

In this mapping system the factor $G_{1\ldots 7}$ appears throughout on the right. It has been introduced to make the system of order 10 with 6-fold points at $P_7{}^4$. If the involution $I_{1\ldots 6}$ is applied to the system this extraneous factor would be reduced to G_7—the directions at p_7—and would disappear from the projective equations of the mapping system.

§ 5

THE GENERALIZED WEDDLE, W_3, IN S_5.

22. Hyperelliptic theta relations ($p=3$) as projective relations in S_5. We set, as in (47),

$$
(78a)\quad
\begin{aligned}
&\Pi=\vartheta\vartheta_{18}\cdots\vartheta_{78},\\
&\pi_{8i_1\ldots i_{2k}}=\vartheta_{i_1\ldots i_{2k}}\cdot\Pi^k/\vartheta\vartheta_{8i_1}\cdots\vartheta_{8i_{2k}},\\
&\pi_{i_1\ldots i_{2k+1}}=\vartheta_{8i_1\ldots i_{2k+1}}\cdot\Pi^k/\vartheta_{8i_1}\cdots\vartheta_{8i_{2k+1}}\\
&\qquad(k=0,\cdots,3;\ i_j=1,\cdots,7).
\end{aligned}
$$

Then, as observed in (61),

(78b) $$\pi_{i_1 \ldots i_{2k+1}} \pi_{8 i_{2k+2} \ldots i_7} = \Pi^2 \cdot \vartheta^2{}_{8 i_1 \ldots i_{2k+1}} = \Pi^2 \cdot \vartheta^2{}_{i_{2k+2} \ldots i_7}.$$

We pass from the functions π to the projective forms by setting

(78c) $$G_{i_1 \ldots i_{2k+1}} = \pi_{i_1 \ldots i_{2k+1}} \cdot \{i_1 \cdot \cdot \cdot i_{2k+1}\}^{1/2};$$

and pass from binary differences $(i_1 i_2)$ to senary determinants in S_5 according to

(78d) $$(i_1 i_2) = \lambda \cdot \epsilon_{i_1 i_2 i_3 \ldots i_8} \quad (i_3 \cdot \cdot \cdot i_8).$$

Since $\vartheta_{i_1 \ldots i_{2k}} = \vartheta_{8 i_{2k+1} \ldots i_7}$ is the significant factor both in $\pi_{8 i_1 \ldots i_{2k}}$ and in $\pi_{i_{2k+1} \ldots i}$, a number of alternative projective forms can be obtained from the same theta relation. Furthermore, in the projective situation, p_8 is no longer isolated, so that relations Bi, given as distinct in **17**, may yield the same projective relation in S_5. Thus we find from the relations of **17** as indicated that:

$$(78e)\quad \begin{array}{rl} \textbf{B3, B1b:} & (14)\,G_{234} \pm (24)\,G_{314} \pm (34)\,G_{124} = 0; \\ \textbf{B1, B2}\ : & G_{238}G_{148} \pm G_{318}G_{248} \pm G_{128}G_{348} = 0; \\ \textbf{B4, B5}\ : & \pm(14)\,(23)\,G_{248}G_{138} \pm (24)\,(13)\,G_{148}G_{238} = G_{12348}; \\ \textbf{B4}\ : & \pm(14)\,(23)\,G_{13567}G_{24567} \pm (24)\,(13)\,G_{14567}G_{23567} = G_{567}G_{1 \ldots 7}. \end{array}$$

In these the binary determinants are to be replaced by senary determinants to secure homogeneity in each point of $P_8{}^5$. The first identity shows that

$$G_{123} = \pm (45678x);$$

the third that G_{12348} is the quadric with nodal plane on p_5, p_6, p_7 and on $p_1, \cdot \cdot \cdot, p_4, p_8$; the fourth that $G_{1 \ldots 7}$ is the cubic spread with triple point at p_8 and double curve N^5; and the second is a projective identity. They all can be obtained by the Clebsch transference principle from corresponding identities in S_4.

The degrees in the points of $P_8{}^5$, and in x, of G_{123}, G_{12345}, and $G_{1 \ldots 7}$ are respectively

$$(0\,0\,0\,1\,1\,1\,1\,1; 1), \quad (2\,2\,2\,2\,2\,4\,4\,4; 2), \quad (6\,6\,6\,6\,6\,6\,6\,9; 3).$$

If we set

(78f)
$$H_{78} = \pm \begin{vmatrix} G_{346}G_{126} & G_{345}G_{125} \\ G_{426}G_{136} & G_{425}G_{135} \end{vmatrix} = \vartheta^2{}_{78}\vartheta^2 \cdot \Pi^2 w \cdot [\{123456\}(78)]^{1/2},$$

$$H_{67} = \pm \begin{vmatrix} G_{348}G_{128} & G_{345}G_{125} \\ G_{428}G_{138} & G_{425}G_{135} \end{vmatrix} = \vartheta^2{}_{68}\vartheta^2{}_{78} \cdot \Pi^2 w \cdot [\{123458\}(67)]^{1/2},$$

where w is defined in (68g), the projective and the analytic expression for H_{78}, H_{67} are reconciled as before by the identity,

$$(78g) \qquad \text{B2b:} \quad H_{78}G_{347}G_{127} \pm H_{68}G_{346}G_{126} \pm H_{58}G_{345}G_{125} \equiv 0.$$

The degrees of H_{78} in P_8^5 and x are (2 2 2 2 2 2 4 4; 4). From (78f) there follows that

$$(78h) \qquad \pm H_{78}H_{56}/(78)(56) = \pm H_{68}H_{75}/(68)(75) = \pm H_{58}H_{67}/(58)(67).$$

In these equations the negative ratios of (78)(56), (68)(75), (58)(67) are to be considered as double ratios of the fundamental octavic. The equations are those of octavic spreads which contain W_3, though we shall find later sextic spreads on W_3.

We need not develop further projective relations as the more important properties of W_3 either have been obtained, or else can be inferred from the properties of its projection W_3'.

23. Geometric properties of W_3. It is comparatively easy to determine the projective equations of the involution I with locus of fixed points, W_3. If $p_1, \cdots, p_6$ are taken as reference points; $G_{178}, \cdots, G_{678}$, as reference S_4's, a provisional form of I is

$$G'_{178} = k_1 G_{23456}/G_{2345678}, \cdots, G'_{678} = k_6 G_{12345}/G_{1234578}.$$

We determine the k_i from the fact that

$$(28)G'_{178} \pm (18)G'_{278} \pm (78)G'_{128} \equiv 0,$$

and $\qquad (28)G_{23456}G_{1345678} \pm (18)G_{13456}G_{2345678} \pm (78)G_{34567}G_{1234568} \equiv 0.$

A comparison of the first two terms shows that $k_1 = \pm k_2$. Hence

(79) *The equations of the involution* $I = I_{1\ldots 8}$ *are*

$$G'_{178} = \pm G_{23456}/G_{2345678}, \cdots, G'_{678} = \pm G_{12345}G_{1234578}.$$

Thus W_3 *is on all the sextic spreads with 4-fold points at* P_8^5 *of the form*

$$G_{178}G_{13456}G_{2345678} \pm G_{278}G_{23456}G_{1345678} = 0.$$

If this spread is recast into theta form, it becomes the identity,

$$\vartheta_{17}\vartheta_{27}\vartheta_{18}\vartheta_{28} - \vartheta_{27}\vartheta_{17}\vartheta_{28}\vartheta_{18} \equiv 0.$$

This double translation of a theta function into G's yields a number of types of sextic spreads on W_3 of which we note the following:

$$
(80) \qquad \begin{aligned}
&G_{234}G_{156}G_{13478}G_{25673} \pm G_{134}G_{256}G_{23478}G_{15678} \equiv 0,\\
&G_{126}G_{678}G_{345}G_{1234578} \pm G_{12678}G_{12345}G_{34578} \equiv 0,\\
&G_{268}G_{167}G_{13458}G_{23457} \pm G_{168}G_{267}G_{23458}G_{13457} \equiv 0.
\end{aligned}
$$

(81) *The variety W_3 has the order* 19. *On it the points of $P_8{}^5$ are 9-fold; the lines p_ip_j and N^5 are 3-fold; and the F-loci of the third kind of $P_8{}^5$, the* 56 *planes $p_i\ p_j\ p_k$, and the* 8 *cones $K_i{}^4$ from respectively p_i to the points of N^5, are simple loci.*

For, W_3 is projected from p_8 into the 10-ic, W_3', with 6-fold points at $P_7{}^4$ and triple lines p_ip_j. The directions at p_8, determined by $\vartheta = 0$, become on W_3' the 2-way cut out by the cubic locus of bisecants of the N^4 on $P_7{}^4$. This cuts W_3' in a 2-way of order 30 with 12-fold points at $P_7{}^4$ from which the seven cubic cones $C_i{}^3$ factor, leaving a 9-ic 2-way with triple points at $P_7{}^4$. Hence p_8 is a 9-fold point on W_3 and the order of W_3 is 19.

W_3 is cut by the P-locus, G_{123}, in a 2-way of order 19 with 9-fold points at $p_4, \cdots, p_8$ from which ten planes, $p_4p_5p_6$, etc. factor leaving a 9-ic 2-way with triple points at $p_4, \cdots, p_8$, and also with triple points at the three points where G_{123} is cut by the lines p_1p_2, p_1p_3, p_2p_3. The complementary P-locus, G_{45678}, must cut W_3 in this same 9-ic 2-way whose transcendental equation is $\vartheta_{1238} = 0$. This quadric, G_{45678}, cuts W_3 in a 38-ic 2-way from which there must separate the plane $p_1p_2p_3$ twice, the 15 planes like $p_1p_2p_4$ each once, and the three cones like $K_1{}^4$ each once, leaving the 9-ic 2-way with triple points at $p_4, \cdots, p_8$.

W_3 is cut by the P-locus, $G_{1\ldots 7}$, in a 57-ic 2-way with 18-fold points at $p_1, \cdots, p_7$ and, normally, a 27-fold point at p_8. This 57-ic consists of the 21 planes like $p_8p_1p_2$, and the seven cones like $K_1{}^4$, and the cone $K_1{}^8$ twice. On this 2-way p_8 has the multiplicity 36 corresponding to the fact that $G_{1\ldots 7}$ contains at p_8 all the directions on W_3 at the 9-fold point p_8.

We observe also that

(82) *W_3 contains the surface $V_2{}^{(5)}$ of* (76), *a 2-way of order* 15 *with* 4*-fold points at $P_8{}^5$.*

We return to the theorem (31a) concerning the linear system Σ of cubic spreads with nodes at $P_8{}^5$. The system Σ contains a linear system Σ_3 consisting of all the members of Σ which contain the F-locus of the second kind, $\pi^{(2)}{}_{1\ldots 8}$, the locus of bisecants of N^5; or, which contain the paired F-locus of the second kind, $N^5 = \pi^{(2)}$, to a multiplicity two, one greater than the normal multiplicity of Σ. The system Σ_3 contains eight members, $G_{2345678}$,

$\cdots, G_{1234567}$, which lie in a one-parameter cubic system with parameters $t = t_1, \cdots, t_8$. The entire system Σ_3 contains the eight F-loci of the third kind, $K_1^4, \cdots, K_8^4$. A member of Σ such as $G_{123} \cdot G_{45678}$ contains three of these only; namely, K_1^4, K_2^4, K_3^4. Also an F-locus of the third kind of type $\pi^{(3)}{}_{123}$ is on three of the above eight members of Σ_3; namely, $G_{2345678}, G_{1345678}$, $G_{1245678}$. In mapping W_3 upon the generalized Kummer 3-way, K_3, in S_7 by means of the system Σ, the $G_{2\ldots 8}$ and $G_{123}G_{4\ldots 8}$ become singular tangent spaces of K_3; and the F-loci of the third kind become the singular points of K_3. Thus $K_1^4, \cdots, K_8^4$ are mapped upon eight points in an S_3 in S_7 which lie upon a cubic curve, the map of the cone K_t^4. Also $G_{23\ldots 8}, \cdots, G_{12\ldots 7}$ are mapped upon eight S_6's containing this S_3, and projecting from this S_3 into eight planes of a cubic curve with parameters $t_1, \cdots, t_8$. Hence (cf. [2] p. 310).

(83) *The Kummer* 3-*way in* S_7 *with* 64 *singular points and* 64 *singular tangent spaces has, in the hyperelliptic case, an additional configuration of* 64 S_3*'s each with the following properties. An* S_3 *is on* 8 *singular points which lie on a cubic space curve with parameters* $t = t_1, \cdots, t_8$; *and is on* 8 *singular spaces which envelop a cubic space curve with parameters* $t = t_1, \cdots, t_8$. *Each of the remaining* 56 *singular points* (*tangent spaces*) *is on three of the* 8 *singular tangent spaces* (*points*).

The existence of this self-dual configuration connected with K_3 is a consequence of the existence of the F-loci of the second kind, intermediate between those of the first and third kinds, in the space of the Weddle W_3.

§ 6

THE HYPERELLIPTIC ($p = 3$) FORM OF THE CAYLEY DIANODE SURFACE.

24. Definition and equation of the surface. A singular space of the Kummer 3-way, K_3^{24}, in S_7 touches K_3 along a 2-way, M_2^{12}. If the functions are not hyperelliptic, the M_2^{12} is the map of Cayley's dianode surface, a sextic surface T with triple points at seven points, P_7^3, in ordinary space, by its linear system of adjoint quartic surfaces [cf. [1] p. 187 (1)]. Schottky has given a parametric equation of the dianodal surface in terms of theta functions of order four for those values u which satisfy $\vartheta(u) = 0$ (cf. [1] p. 181). However, the situation is quite different in the hyperelliptic case. Since we hope to make in the future a closer study of the hyperelliptic Kummer manifold K_p with the aid of the mappings indicated in this article, it seems pertinent to develop here the peculiarities of Schottky's representation in the hyperelliptic case.

The representation is based on the following theta relations of types A and C [cf. (**17**)]:

$$\vartheta(u)=0;$$

$$\begin{aligned}(84\text{a})\quad & [\{238\}\{156\}]^{1/2}\,\vartheta_{1478}\vartheta_{18}\pm[\{318\}\{256\}]^{1/2}\,\vartheta_{2478}\vartheta_{28}\\ &\qquad\pm[\{128\}\{356\}]^{1/2}\vartheta_{3478}\vartheta_{38}=0;\\ & [(12)(34)(58)(67)]^{1/2}\vartheta_{23}\vartheta_{14}\pm[(14)(23)(57)(68)]^{1/2}\vartheta_{1278}\vartheta_{3478}\\ &\qquad\pm[(14)(23)(56)(78)]^{1/2}\,\vartheta_{1268}\vartheta_{3468}=0;\\ & [\{234\}\{856\}]^{1/2}\,\vartheta_{1478}\vartheta_{48}\pm[\{438\}\{256\}]^{1/2}\,\vartheta_{1247}\vartheta_{24}\\ &\qquad\pm[\{428\}\{356\}]^{1/2}\,\vartheta_{1347}\vartheta_{34}=0.\end{aligned}$$

The functions introduced by Schottky are

$$\begin{aligned}(84\text{b})\qquad \Pi &= \vartheta_{18}\vartheta_{28}\cdot\cdot\cdot\vartheta_{78};\\ P_{ijk} &= \vartheta_{ijk8}\vartheta_{i8}\vartheta_{j8}\vartheta_{k8};\\ P_{i,j} &= \vartheta_{ij}\cdot\Pi\vartheta_{i8}/\vartheta_{j8};\\ P_{ijkl} &= \vartheta_{mno8}\cdot\Pi^2/\vartheta_{m8}\vartheta_{n8}\vartheta_{o8};\\ P_i &= \vartheta^2{}_{i8}\cdot\Pi^2\\ & (i,j,\cdot\cdot\cdot,o=1,\cdot\cdot\cdot,7).\end{aligned}$$

We convert some of these into projective forms by setting

$$\begin{aligned}(84\text{c})\qquad G_{123} &= \pm P_{123}\cdot\{123\}/[\{8;123\}\{4567\}]^{1/2},\\ G_{1,2} &= \pm P_{1,2}\cdot\{134567\}\{34567\}/[\{234567\}\{8;134567\}(23)]^{1/2}.\end{aligned}$$

The first relation (84a) then yields

$$\{2347\}G_{147}\pm\{3147\}G_{247}\pm\{1247\}G_{347}\equiv 0.$$

This shows, in the customary way, that the functions G_{123} are linear sections of a theta manifold in S_3 by planes on the points p_1, p_2, p_3 respectively of a set $P_7{}^3$. Moreover the coefficients of the relations are such that the $P_7{}^3$ is on a norm-curve N^3 with parameters $t_1,\cdot\cdot\cdot,t_7$. If N^3 is taken as

$$\begin{aligned}(84\text{d})\qquad & x_0=t^3,\quad x_1=t^2,\quad x_2=t,\quad x_3=1,\qquad \text{then}\\ & G_{147}=\pm(147x),\quad \{2347\}=\pm(2347).\end{aligned}$$

With this transition to projective quaternary determinants the second relation (84a) becomes

$$\begin{aligned}(84\text{e})\qquad G_{3,2} &= G_{356}G_{347}(1357)(1346)-G_{357}G_{346}(1356)(1347)\\ &=(13)^2\{13;4567\}[G_{356}G_{347}(57)(46)-G_{357}G_{346}(56)(47)].\end{aligned}$$

The first expression for $G_{3,2}$ is purely projective, and shows that $G_{3,2}$ is a quadric with node at p_3 and simple points at p_4, p_5, p_6, p_7. From the symmetry of its definition in (84c), the quadric $G_{3,2}$ must pass through p_1 also.

But since in the present case P_7^3 is on N^3, $G_{3,2}$ must contain N^3, and therefore also the point p_2. Thus it must differ from $G_{3,1}$ by a constant factor only, and we find at once from the second form of $G_{3,2}$ in (84e) that

$$(84f)\qquad (2357)(2346)G_{3,2} \pm (1357)(1346)G_{3,1} \equiv 0.$$

If this projective identity is recast into transcendental form it becomes

$$(84g)\qquad [(82)(31)]^{1/2}\vartheta_{23}\vartheta_{18} \pm [(81)(23)]^{1/2}\,\vartheta_{31}\vartheta_{28} \equiv 0,$$

which is the form the identity B4b in **17** takes when $\vartheta(u)=0$.

In order to give $G_{3,2}$ in (84e) a specific sign we shall write

$$(84h)\qquad \begin{aligned} G_{147} &= (147x) \\ G_{3,2} &= W(3;\ 14567x), \end{aligned}$$

and agree that corresponding functions shall arise from the latter by even permutations of the arguments $1,\cdots,7,x$. Thus $W(x;\ 123456) = \pm\, G_{x,7}$ is the Weddle quartic surface with nodes at $p_1,\cdots,p_6$, and is to be taken with the specific sign which arises when the permutation, $(27)(3x)$, is applied to $G_{3,2}$ in (84e). The function W is an alternating function of its last six arguments.

The third of the relations (84a) can be recast into

$$(84i)\qquad \begin{aligned} &P_{2356}\cdot\{8;17\}\{2317\}\{2356\}\{235617\}/[\{12345678\}\{147\}]^{1/2} \\ &\qquad\equiv \pm(8137)G_{356}G_{2,4} \pm (8127)G_{256}G_{3,4}. \end{aligned}$$

The right member is purely projective. Hence P_{2356} is a cubic surface with nodes at p_2, p_3, p_5, p_6 and on p_1, p_7; but also on p_4 since $G_{2,4}$, $G_{3,4}$ are also on p_4. When P_7^3 is on N^3 there is a pencil of such surfaces (the image of a pencil of planes under a regular cubic transformation). If p_8 is taken on N^3 with parameter $t=t_8$, the binary factors (81) and (82) separate from (8137) and (8127) leaving t_8 as the linear parameter of the pencil.

The identical relations ([1] p. 183) among the functions now lead in part to projective identities. For example, the first one, $P_{1,2}P_{2,3}P_{3,1} = P_{2,1}P_{1,3}P_{3,2}$, reduces by the use of (84b, c) to $G_{1,2}G_{2,3}G_{3,1} = \pm\, G_{2,1}G_{1,3}G_{3,2}$, which, according to (82f), is an identity. But the second relation,

$$(85a)\qquad P_{3,4}P_{124}P_{4567} - P_{4,3}P_{123}P_{3567} \equiv 0,$$

is converted by the use of (84c, e, i) into the projective relation,

$$(85b)\qquad \begin{aligned} S \equiv\ & G_{3,4}G_{124}\,[(1278)G_{457}G_{6,3} \mp (1268)G_{456}G_{7,3}] \\ & \mp G_{4,3}G_{123}\,[(1278)G_{357}G_{6,4} \mp (1268)G_{356}G_{7,4}] \equiv 0. \end{aligned}$$

This is a relation of the sixth order, an identity in u for values subject to $\vartheta(u)=0$, but not an identity in x for all positions of p_8 on N^3; and therefore

is the equation of the sextic surface whose parametric equations are furnished by (84c). For variable $p_8 = t_8$ on N^3, the factors (81), (82) separate from (1278) and (1268), so that the sextic surface is a member of a pencil with linear parameter t_8. The degrees of each term of the equation in $p_1, \cdots, p_8, x$ are 6, 6, 4, 4, 5, 6, 6, 1, 6.

On isolating the parameter, (1278) : (1268), of the pencil it reads:

$$S = (1278)\,[G_{3,4}G_{6,3}G_{124}G_{457} \mp G_{4,3}G_{6,4}G_{123}G_{357}] \mp (1268)\,[G_{3,4}G_{7,3}G_{124}G_{456} \mp G_{4,3}G_{7,4}G_{123}G_{356}] = 0. \tag{85c}$$

Each term is quadratic in $G_{i,j}$ and therefore the pencil contains N^3 doubly. Each term contains $G_{4,3}$ or $G_{3,4}$ and therefore the pencil contains the line p_3p_4. It must therefore contain each of the 21 lines p_ip_j. Each of the two terms of the first member of the pencil contains explicitly all of the 21 lines which are on p_6, p_3, p_4 as well as the lines p_1p_2, p_5p_7, leaving for examination only the lines from p_1, p_2 to p_5, p_7. With the definition of sign in (84e, h) we find, when x is taken to be $\lambda_1p_1 + \lambda_2p_5$, that the term in $\lambda_1{}^3\lambda_2{}^3$ cancels in (85c) only when the ambiguous sign within the brackets is negative, or when the ambiguous sign outside the brackets in (85b) is negative.

The same determination of sign permits of writing the more specific form of (84f):

$$(2357)\,(2346)\,G_{3,2} + (1357)\,(1346)\,G_{3,1} \equiv 0. \tag{85d}$$

We use this relation to remove the factor $G_{6,4} = kG_{6,3}$ from the first member of the pencil (85c). It becomes, to within the factor $-(1278)\,G_{6,4}/(6312)\,(6357)$,

$$\{(6412)\,(6457)\,G_{3,4}G_{124}G_{457} + (6312)\,(6357)\,G_{4,3}G_{123}G_{357}\}.$$

We prove that this brace is a Weddle surface as follows: replace $G_{3,4}$ and $G_{4,3}$ in the brace by the values,

$$G_{3,4} = (6357)\,(6312)\,G_{351}G_{327} - (6351)\,(6327)\,G_{357}G_{321},$$
$$G_{4,3} = (6457)\,(6412)\,G_{452}G_{417} - (6452)\,(6417)\,G_{457}G_{412}.$$

If now the quaternary determinants are all expressed in terms of the binary differences the factor

$$m = \{6;\ 123457\}^2\{34;\ 1257\}\,(12)\,(57)$$

can be removed. The two like products of four G's then have further coefficients (51)(27), (25)(17) which coalesce into (57)(21). The brace then becomes

$$m \cdot (12)\,(57) \cdot [-G_{124}G_{574}G_{357}G_{312} - G_{124}G_{574}G_{351}G_{327} + G_{123}G_{573}G_{174}G_{524}].$$

The first two terms in the bracket coalesce into $G_{124}G_{574}G_{352}G_{871}$, and the bracket is therefore the Weddle quartic surface $G_{x,6}$ with nodes at the points of P_7^3 other than p_6. Collecting all the terms, and passing to the second member of the pencil (85c) by interchanging 6,7 we find that

(85e) $S=(4567)[(1278)(1246)G_{6,4}G_{x,6}\mp(1268)(1247)G_{7,4}G_{x,7}]$.

The facts concerning the surface S, and the pencil in which it lies, may now be summarized as follows:

(86) *The pencil of sextic surfaces S has triple points at P_7^3, has the double curve N^3 asymptotic on one sheet of the surface, and contains the 21 lines p_ip_j. As p_8 with parameter $t=t_8$ runs along N^3, the surface S for $t=t_i$ $(i=1,\cdots,7)$ degenerates into $G_{i,j}G_{x,i}$. The equation of the surface is obtained from any determinant of this (or similar) matrix,*

$$\left|\begin{matrix} (45)(67)G_{5,4}G_{x,5} & (46)(75)G_{6,4}G_{x,6} & (47)(56)G_{7,4}G_{x,7} \\ (85)(67) & (86)(75) & (87)(56) \end{matrix}\right| = 0,$$

in which the binary differences are to be interpreted as terms of double ratios. With the determinations of sign in (84e, h) *the upper signs are to be chosen in* (85b, c, e).

For, we see in (85e) that when $p_8=p_6$, S becomes $G_{6,4}G_{x,6}$. The cone $G_{6,4}$ contains N^3, and the Weddle surface $G_{x,6}$ contains N^3 as an asymptotic curve. We shall verify in the next section that the sum of the elements in the upper row of the given matrix is identically zero [cf. (87e)], and evidently this is true of the lower row also. The last determinant of the matrix reduces to (85e) with upper sign. Since S also admits similar representations in terms of the other determinants, the three representations can coexist only for the given choice of signs.

We observe that the base curve of order 36 of the pencil consists of N^3 five-fold, and the 21 lines p_ip_j.

25. The surface S in a binary notation with reference to N^3. We use that coördinate system in S_3 in which the coördinates of a point are the coefficients of a binary cubic and perfect cubes are points on N^3. The notation is that of Gordan ([13] Part II, pp. 172-3) in which

$$\text{(87a)}\quad \begin{aligned} & f=(a_1t)^3=(a_2t)^3=\cdots; \\ & \Delta=(a_1a_2)^2(a_1t)(a_2t)=(\Delta t)^2=(\Delta_1 t)^2=\cdots; \\ & Q=(a_1a_2)^2(a_3a_2)(a_1t)(a_3t)^2=(Qt)^3=(Q_1t)^3=\cdots; \\ & R=(\Delta\Delta_1)^2; \\ & 2Q^2+Rf^2+\Delta^3\equiv 0;\ \text{and let also} \\ & k=f\cdot Q=(kt)^6=(k_1t)^6=\cdots. \end{aligned}$$

The quaternary determinants then reduce to binary products as in

(87b) $$G_{123} = \{123\} \cdot (a_1 t_1)(a_1 t_2)(a_1 t_3)\,;\ (1234) = \{1234\}.$$

The cone $G_{3,2}$ in (84e), and the Weddle surface $G_{x,3}$, are now

(87c) $$\begin{aligned} G_{3,2} &= -\{3\,;\,14567\}^2\{14567\} \cdot (\Delta t_3)^2/2, \\ G_{x,3} &= c \cdot \{124567\} \cdot (k\alpha_3)^6, \end{aligned}$$

where the numerical constant c need not be evaluated, and where $(\alpha_3 t)^6$ is the sextic with roots $t_1, \cdots, t_7$ excepting t_3.

We first prove that

(87d) $$\begin{aligned} (67) \cdot (\Delta t_5)^2 \cdot (kt_6)(kt_7)(kt)^4 + (75) \cdot (\Delta t_6)^2 \cdot (kt_7)(kt_5)(kt)^4 \\ + (56) \cdot (\Delta t_7)^2 \cdot (kt_5)(kt_6)(kt)^4 \equiv 0. \end{aligned}$$

For, the left member changes sign when any two of t_5, t_6, t_7 are interchanged, and therefore contains the factor $(56)(57)(67)$. The residual factor is a covariant of the cubic, of order 4 and degree 6, which cannot exist.

If the identity be polarized with respect to t_1, t_2, t_3, t_4, it becomes by virtue of (87c),

(87e) $$(45)(67)G_{5,4}G_{x,5} + (46)(75)G_{6,4}G_{x,6} + (47)(56)G_{7,4}G_{x,7} \equiv 0.$$

We now prove the theorem:

(88a) *The pencil of sextic surfaces (86) may be defined as that pencil with triple points at $P_7{}^3$ only on N^3, whose members are individually invariant under the cubic Cremona involution J whose generic pair x, x' is on a bisecant of N^3 and harmonic with the intersections of the bisecant.*

Under J the points represented by f and Q are interchanged and then ([13] § 149)

$$f' = Q, \quad \Delta' = R\Delta/2, \quad Q' = -Rf^2/4, \quad R' = R^3/4.$$

The P-surface of J is the quartic surface, $R = 0$, the locus of tangents of N^3, on which N^3 is a double doubly-asymptotic curve (cf. [14] p. 4). If J is to leave a sextic surface unaltered, it must transform the surface into itself multiplied by R^3. The only independent covariants of f of degree 6 which acquire a factor R^3 are

$$f\Delta Q, \quad Rf^2, \quad R\Delta, \quad Q^2$$

which are transformed to within the factor $R^3/8$ into respectively

$$-f\Delta Q, \quad 2Q^2, \quad R\Delta, \quad Rf^2/2.$$

Hence $$R\Delta, \quad Rf^2 + 2Q^2, \quad Rf^2 - 2Q^2, \quad f\Delta Q$$

are transformed respectively into

$$R\Delta, \quad Rf^2+2Q^2, \quad -(Rf^2-2Q^2), \quad -f\Delta Q.$$

Hence, replacing Rf^2 by using the syzygy (87a),

(88b) *Under J there are two systems of invariant sextic surfaces, namely,*

$$(R\Delta, \alpha)^2+(\Delta^3, \beta)^6=0,$$
$$(f\Delta Q, \gamma)^8+(4Q^2+\Delta^3, \delta)^6=0,$$

where the coefficients of $(\alpha t)^2$, $(\beta t)^6$, $(\gamma t)^8$, $(\delta t)^6$ in the indicated apolarity invariants are the parameters of the systems.

The coefficients of Δ contain N^3 simply, those of Q contain N^3 as a simple simply-asymptotic curve. Hence the first system contains N^3 as a triple curve. The second system, however, contains N^3 as a double simply-asymptotic curve. We wish to find the conditions on $(\gamma t)^8$, $(\delta t)^6$ that a surface of this second system may have a triple point at $t=t_1$. The part containing Δ^3, which has the triple curve N^3, may be dropped, and we have to examine

$$(a_1\gamma)^3(a_2a_3)^2(a_2\gamma)(a_3\gamma)(a_4a_5)^2(a_6a_5)(a_4\gamma)(a_6\gamma)^2$$
$$+4(a_1a_2)^2(a_3a_2)(a_1\delta)(a_3\delta)^2(a_4a_5)^2(a_6a_5)(a_4\delta)(a_6\delta)^2=0.$$

The quadric polar of the point (t_1t) is obtained by setting four of the six symbols, a_{i0}, a_{i1} (in each of 15 ways) equal to t_{11}, $-t_{10}$, leaving the other two as variable, and dividing by 15. The result is

$$2[(\gamma t_1)^7(a_1\gamma)(a_1t_1)^2+2(a_1t_1)^3\cdot(\delta t_1)^6]\cdot(a_2t_1)^3/15.$$

The factor, $(a_2t_1)^3$, the osculating plane of N^3 at t_1, accounts for the fact that N^3 is a simply-asymptotic curve. The vanishing of the bracket is the significant part of the condition on $(\gamma t)^8$, $(\delta t)^6$ that the surface may have a triple point at t_1. On replacing symbols a_1 by t and factoring out $(t_1t)^2$, this condition on γ, δ reduces to the identity, linear in t,

$$(\gamma t_1)^7(\gamma t)+2(\delta t_1)^6\cdot(t_1t)\equiv 0.$$

On setting $t=t_1$ in this identity, $(\gamma t_1)^8=0$.

Thus to secure triple points at P_7^3 we set

$$(88c) \qquad \begin{aligned}(\sigma t)^7&=(t_1t)\cdots(t_7t),\\ (\gamma t)^8&=(t_8t)\cdot(\sigma t)^7,\end{aligned}$$

the t_8 remaining in γ as a linear parameter. Then

$$8(\gamma t_1)^7(\gamma t)=\{2345678;1\}\cdot(t_1t).$$

In order to satisfy the above linear identity at P_7^3, the coefficients of $(\delta t)^6$ must satisfy the seven linear non-homogeneous equations,

$$16(\delta t_1)^6 = \{1;\ 2345678\}, \cdots, 16(\delta t_7)^6 = \{7;\ 1234568\}.$$

Now $7(\sigma t)^6(\sigma t_8) = \Sigma(18)(t_2 t) \cdots (t_7 t)$, and $7(\sigma t_1)^6(\sigma t_8) = \{1;\ 2345678\} = 16(\delta t_1)^6$, etc. Hence $(\delta t)^6 = 7(\sigma t_8)(\sigma t)^6/16$.

(88d) *With* $(\gamma t)^8$ *and* $(\sigma t)^7$ *defined as in* (88c) *the pencil* (*for variable* t_8) *of sextic surfaces,*

$$16\,[f\Delta Q, \gamma]^8 + 7\,[4Q^2 + \Delta^3, (\sigma t_8)(\sigma t)^6] = 0,$$

has triple points at $P_7{}^3$ *only on* N^3; *and each member is invariant under* J.

In order to complete the proof of (88a) it is necessary to show that the member of the pencil (88d) obtained by setting $t_8 = t_1$ is the product of the quadric cone on N^3 with vertex at p_1, and the Weddle with nodes at $p_2, \cdots, p_7$, i. e., is

$$(\Delta t_1)^2 \cdot (fQ, \alpha_1)^6 = (\Delta t_1)^2 \cdot (k\alpha_1)^6 \qquad [\text{cf. } (87c)].$$

When $t_8 = t_1$, $(\gamma t)^8 = (t_1 t)^2 \cdot (\alpha_1 t)^6$, $(\sigma t_1)(\sigma t)^6 = (6(t_1 t) \cdot (\alpha_1 t_1)(\alpha_1 t)^5/7$. The member of the pencil (88d) is then

$$\begin{aligned} &8[\Delta k,\ (t_1 t)^2 \cdot (\alpha_1 t)^6]^8 + 3[4Q^2 + \Delta^3,\ (t_1 t) \cdot (\alpha_1 t_1)(\alpha_1 t)^5]^6 \\ &\quad = 8[\{\Delta k\}_{t_1^2}, (\alpha_1 t)^6]^6 + 3[\{4Q^2 + \Delta^3\}_{t_1}, (\alpha_1 t_1)(\alpha_1 t)^5]^5 = 0, \end{aligned}$$

where subscripts t_1^2, t_1 indicate polarizations. The second polar form can be expanded as follows:

$$\begin{aligned} \{\Delta k\}_{t_1^2} = (\Delta t_1)^2 \cdot (kt)^6 - 3(\Delta k)(\Delta t)(kt_1)(kt)^4 \cdot (t_1 t)/2 \\ + 15(\Delta k)^2 (kt)^4 \cdot (t_1 t)^2/28. \end{aligned}$$

The first term of this expansion contributes the desired product,

$$8(\Delta t_1)^2 \cdot (k\alpha_1)^6.$$

The last term, involving a covariant, $(\Delta k)^2(kt)^4$, of order 4 and degree 6, vanishes identically. Since the last term does vanish identically, the middle term can be written as $-3\{(\Delta k)(\Delta t)(kt)^5\}_{t_1} \cdot (t_1 t)/2$. But the transvectant

$$\begin{aligned} (\Delta k)(\Delta t)(kt)^5 &= (\Delta, fQ)^1 \\ &= [f(\Delta, Q)^1 + Q(\Delta, f)^1]/2 = -(4Q^2 + \Delta^3)/4 \end{aligned}$$

(cf. [13]). Hence the member can be written as

$$\begin{aligned} 8(\Delta t_1)^2 \cdot (k\alpha_1)^6 + 3[\{4Q^2 + \Delta^3\}_{t_1} \cdot (t_1 t), (\alpha_1 t)^6]^6 \\ + 3[\{4Q^2 + \Delta^3\}_{t_1}, (\alpha_1 t_1)(\alpha_1 t)^5]^5 = 0. \end{aligned}$$

In this the last two transvectants cancel and the proof is complete.

We have not yet determined the geometric character of the point $t_8 = p_8$ on the member of the pencil which it determines. This is as follows:

(89) *Each member of the pencil of sextic surfaces S which is not degenerate has a pinch point at one point $p_8 = t_8$ on its double curve.*

For, if, in the quadric polar of $(t_1t)^3$ above, we replace t_1 by t_8, and omit the osculating plane of N^3, $2(a_2t_8)^3/15$, the remaining factor,

$$(\gamma t_8)^7(a_1\gamma)(a_1t_8)^2 + 2(a_1t_8)^3 \cdot (\delta t_8)^6,$$

becomes, due to the above determinations of γ and δ, precisely

$$\{1234567;\, 8\} \cdot (a_1t_8)^3.$$

This also is the osculating plane of N^3 except at the triple points

$$t_8 = t_1, \cdot \cdot \cdot, t_8 = t_7.$$

Furthermore this part of the quadric polar of the generic point t of N^3, $(\gamma t)^7(a_1\gamma)(a_1t)^2 + 2(a_1t)^3 \cdot (\delta t)^6 = 0$, reduces to $(a_1t_8)(a_1t)^2 \cdot (\sigma t)^7 = 0$ which is the plane on p_8 and tangent to N^3 at t, i. e., the tangent plane at t of the quadric cone, G_8, on N^3 with vertex at p_8. Hence

(90 *The surface S determined by $t_8 = p_8$ has a sheet passing through N^3 and tangent along N^3 to the quadric cone G_8 with vertex at p_8.*

26. Mapping of S upon the section of K_3 by a singular space. From (84b) there follows that

$$\begin{aligned} \Pi^2 \cdot \vartheta^2{}_{ij} &= P_{i,j}P_{j,i}, \\ \Pi^2 \cdot \vartheta^2{}_{ijk8} &= P_{ijk}P_{lmno}, \qquad (91) \\ \Pi^2 \cdot \vartheta^2{}_{i8} &= P_i \\ & \qquad (i, j, \cdot \cdot \cdot = 1, \cdot \cdot \cdot, 7). \end{aligned}$$

The products $P_{i,j}P_{ji}$ and $P_{ijk}P_{lmno}$ have been converted into quartic surfaces on N^3 with nodes at $P_7{}^3$. We identify P_i with a similar quartic surface in precisely the same way as in ([1] p. 183). With $i = 7$, P_7 is proportional to

$$\{238\}(16)(17)G_{1,6}G_{7,1} \pm \{318\}(26)(27)G_{2,6}G_{7,2} \pm \{128\}(36)(37)G_{3,6}G_{7,3}.$$

If from this we eliminate $G_{7,2}$ and $G_{7,3}$ by using [cf. 84(f)]

$$\begin{aligned} (1734)(1756)G_{7,1} &= \pm(2734)(2756)G_{7,2}, \\ (1724)(1756)G_{7,1} &= \pm(3724)(3756)G_{7,3}, \end{aligned}$$

there results

$$G_{7,1} \cdot [\{234578\}(23)G_{1,6} \pm \{314578\}(31)G_{2,6} \pm \{124578\}(12)G_{3,6}] = 0.$$

On inserting the binary notation from (87c), this becomes

$$(\Delta t_7)^2 \cdot [\{238\} \cdot (\Delta t_1)^2 \pm \{318\} \cdot (\Delta t_2)^2 \pm \{128\} \cdot (\Delta t_3)^2] = 0.$$

Since from its definition P_7 must be symmetrical with respect to $p_1, \cdot \cdot \cdot, p_6$,

the $\pm$ signs must be such that the second factor of P_7 must be $\pm\{123\}\cdot(\Delta t_8)^2$ whence

(92) $$P_7 = k\cdot(\Delta t_7)^2\cdot(\Delta t_8)^2.$$

Thus P_7, the product of the two quadric cones, G_7, G_8 on N^3 with vertices at p_7 and p_8 respectively, is in the linear system of quartic surfaces which contain N^3 and have nodes at $P_7{}^3$.

Quartic surfaces of this system are found in the system which contains N^3, which is made up as follows:

(93) $$(\Delta f^2, \gamma)^8 + (fQ, \delta)^6 + (\Delta^2, \epsilon)^4 + \zeta R = 0.$$

The last two terms comprise those surfaces which contain N^3 doubly so that we seek to determine the forms $(\gamma t)^8$, $(\delta t)^6$ so that the system may have double points at $t_1, \cdots, t_7$ on N^3. The polar plane of the point represented by $(t_1 t)^3$ is $(at_1)^3\cdot(\delta t_1)^6 + 2(at_1)^2(a\gamma)(\gamma t_1)^7 = 0$. If this is to vanish for all points, $(at)^3$, it must vanish identically when symbols a are replaced by variables t from $(at') = (tt')$; i. e.,

$$(tt_1)\cdot(\delta t_1)^6 - 2(\gamma t)(\gamma t_1)^7 \equiv 0.$$

For the particular value $t = t_1$, $(\gamma t_1)^8 = 0$, whence we introduce γ and σ as in (88c), and find as before that $4(\delta t)^6 = 7(\sigma t_8)(\sigma t)^6$. With these values of γ and δ, the polar plane of a general point t on N^3 has the equation, $(at)^2(at_8)\cdot(\sigma t)^7 = 0$. If t is not at a point of $P_7{}^3$, this is the plane on t_8 containing the tangent at t, i. e., is the tangent plane of G_8 at t; whereas if t is at t_8, it is the osculating plane of N^3 at t_8. Hence

(94) *The linear system of theta squares, other than $\vartheta^2(u) = 0$, are represented on S by the linear system of ∞^6 quartic surfaces,*

$$4[\Delta f^2, (t_8 t)\cdot(\sigma t)^7]^8 + 7[fQ, (\sigma t_8)(\sigma t)^6]^6 + (\Delta^2, \epsilon)^4 + \zeta R = 0,$$

where $(\sigma t)^7 = (t_1 t)\cdots(t_7 t)$, and where ζ and the coefficients of $(\epsilon t)^4$ are the six non-homogeneous parameters of the system. This system contains N^3, touches G_8 along N^3, and at $t_8 = p_8$ has for tangent plane the osculating plane of N^3. The system maps S upon the section of the Kummer 3-way, K_3, in S_7 by its singular tangent space, $\vartheta(u) = 0$.

A projection of this section of K_3 from three of the singular points of K_3 upon the section, which is a sextic surface in ordinary space, is the subject of an elegant memoir by Humbert.[15] This projection is thus a birational transform of the Cayley dianodal surface. It has of course a much lower degree of symmetry with respect to the theta functions than the Cayley surface. A further development of the properties of the surface of Humbert, in both the general and the hyperelliptic case, is given by L. Remy.[16, 17]

URBANA, ILLINOIS.

REFERENCES.

[1] A. B. Coble, "Algebraic Geometry and Theta Functions," *Colloquium Publications of the American Mathematical Society*, New York, Vol. 10 (1929).

[2] A. B. Coble, "The Equation of the Eighth Degree," *Bulletin of the American Mathematical Society*, Vol. 30 (1924), pp. 301-313.

[3] Report, "Selected Topics in Algebraic Geometry," *Bulletin, National Research Council*, No. 63 (April, 1928). Washington.

[4] A. B. Coble, "Point Sets and Allied Cremona Groups," *Transactions of the American Mathematical Society*, I: Vol. 16 (1915), pp. 155-198; II: Vol. 17 (1916), pp. 345-385; III: Vol. 18 (1917), pp. 331-372.

[5] F. Schottky, "Ueber die Beziehungen zwischen den sechzehn Theta-functionen von zwei Variabeln," *Journal für Mathematik*, Vol. 105 (1889), pp. 233-249.

[6] A. B. Coble, "The Ten Nodes of the Rational Sextic and of the Cayley Symmetroid," *American Journal of Mathematics*, Vol 41 (1919), pp. 243-265.

[7] A. B. Coble, "Associated Sets of Points," *Transactions of the American Mathematical Society*, Vol. 24 (1922), pp. 1-20.

[8] A. Krazer, *Lehrbuch der Thetafunktionen*, Teubner, Leipzig (1903).

[9] J. Thomae, "Beitrag zur Bestimmung von θ $(0, 0, \cdots, 0)$ durch die Klassenmoduln algebraischer Functionen," *Journal für Mathematik*, Vol. 71 (1870), pp. 201-222.

[10] O. Zariski, "On Hyperelliptic Theta-functions with Rational Characteristics," *American Journal of Mathematics*, Vol. 50 (1928), pp. 315-344.

[11] H. F. Baker, "Elementary Note on the Weddle Quartic Surface," *Proceedings of the London Mathematical Society* (2), Vol. 1 (1904), pp. 247-261.

[12] R. W. H. T. Hudson, *Kummer's Quartic Surface*, Cambridge (1905).

[13] P. Gordan, *Invarianten-theorie*, Teubner, Leipzig (1885).

[14] A. B. Coble, "On the Relation between the Three-parameter Groups of a Cubic Space Curve and a Quadric Surface," *Transactions of the American Mathematical Society*, Vol. 6 (1906), pp. 1-20.

[15] G. Humbert, "Sur une surface du sixième ordre liée aux fonctions abéliennes de genre trois," *Journal de Mathématiques*, Vol. 61 (1896), pp. 263-293.

[16] L. Remy, "Sur certaines surfaces algébriques liées aux fonctions abéliennes de genre trois," *Journal de Mathématiques*, Vol. 73 (1908), pp. 1-37.

[17] L. Remy, "Sur une classe de surfaces algébriques liées aux fonctions abéliennes de genre trois," *Annales de l'École Normale Supérieure*, Vol. 45 (1909), pp. 193-274.

On the Moduli of Algebraic Functions.

By Francesco Severi.

(Extract from a letter to Professor Oscar Zariski.)

In your paper, published in this Journal, January 1930, pp. 150-170, you observe that the reasoning employed in sec. 5 of my article "Sul Teorema di Esistenza di Riemann," *Rendiconti del Circolo Matematico di Palermo*, (1922), proves that on a curve Γ of genus $p(>1)$, of general moduli, the two-fold of a *generic* g_ν^1, of any order ν (necessarily $\geqq p/2+1$), is non-special; but that the question remains whether this holds for *any* g_ν^1 of Γ.

I wish to show you that the above theorem is also valid if referred to *any* g_ν^1 of Γ, and that the essential elements necessary to reach this conclusion are already contained in my quoted article and in my "Vorlesungen über algebraische Geometrie," to which that article refers. Your remark is at any rate opportune, since perhaps the deduction was too concise, and therefore it will be useful to develop it with more details, as I now intend to do. In the sequel I use the notations of my article, and I explicitly remind you that the curves, which are considered below, are always supposed to be of genus $p>1$.

First of all a remark of a general character. Let H be an arbitrary irreducible system, ∞^r, of curves C of genus p, and let H contain $\infty^{r'}(r'<r)$ curves C_0, constituting a system H' and enjoying a given property, which is not enjoyed by a generic C (a priori this property need not be invariant under birational transformations). Let the curves C birationally equivalent to a generic C be ∞^δ, and let the curves C_0 of H', which are birationally equivalent to a generic C_0, be $\infty^{\delta'}$. If $\delta'\geqq\delta$ (this is the essential hypothesis!), then denoting by μ, μ' the number of moduli of the variable curve C of H and of the variable curve C_0 of H' respectively, we have evidently the following relations:

$$\mu=r-\delta,\qquad \mu'=r'-\delta'<\mu-(\delta'-\delta),$$

and hence the C_0's, as curves of H, are of special moduli.

In our case, H is the irreducible system of the curves C, of order $n+\nu$ and genus p, having two multiple points of orders ν, n at P, Q respectively *

* The integer n must satisfy the only condition that series g_n^1 exist on every curve of genus p. Accordingly we must have $n\geqq p/2+1$.

and $d=(n-1)(\nu-1)-p$ distinct ordinary double points; H' is the system of those curves C_0 of H, which enjoy the particular property, that the index of speciality i' of the two-fold of the $g_\nu{}^1$, cut out on a generic C_0 by the lines of the pencil Q, is *greater* than the index of speciality $i(\geqq 0)$ of the two-fold of the similar $g_\nu{}^1$, relative to a generic C; $r=2n+2\nu+p-1$; $\delta=2n+2\nu-2p+2+i$. As far as δ' is concerned, I shall show later that

$$(1) \qquad \delta' \geqq 2n+2\nu-2p+2+i',$$

whence it follows that

$$\mu' < \mu-(\delta'-\delta) \leqq 3p-3-i'.$$

Hence, in the case of curves of general moduli ($i=0$), the following theorem holds:

On a curve of general moduli of genus p, the two-fold of any *infinite linear series is non-special.*

For $i>0$, the same reasoning, which leads to (1), with a few additional considerations, serves to prove that:

On the most general curve of genus p, containing a $g_\nu{}^1$, $\nu < p/2+1$, without fixed points, the two-fold of any *series $g_\nu{}^1$, without fixed points, has its index of speciality $i>0$, $3p-3-i$ being the number of moduli on which that curve depends.*

Moreover it is known, as was shown by my assistant, Professor B. Segre,* and as you yourself found out again in your paper, that the number i for the above curve is $i=p-2(\nu-1)$.

You will observe that the above theorems also serve to reveal in general the exact significance of your remarks concerning a quite secondary and incidental assertion contained in the quoted paper by B. Segre. It is the same author, to whom I have communicated the present considerations, who has called my attention to it. In fact, if $i'=p-2(\nu-1)+\sigma$, with $\sigma>0$, the number of moduli on which the curve depends falls *below* $2p+2\nu-5-\sigma$, in agreement with what you have verified in an example ($\sigma=1$) by a detailed and a delicate analysis.

To derive (1), I shall prove in the first place that:

1) Given an *arbitrary* curve Γ of genus p and assigned *arbitrarily* a

---

* *Mathematische Annalen*, Vol. 100 (1928).

g_ν^1, without fixed points, on Γ and an integer $n > 2p + 1$; if furthermore H is the system, relative to the above values of n, ν, p, then the curve C_0 of H, birationally equivalent to Γ, which is being constructed by starting from the above g_ν^1 and from a *generic* g_n^1 on Γ and by setting up a projective correspondence between the groups of these series and the lines of the pencils Q, P, has the same point singularities as the generic curve C of H.

2) The $\sigma = 2n + 2p - 2$ tangents λ, drawn from P and touching C_0 outside P, have, as in the case of a generic C, a two-point contact with C_0 (and hence are distinct).

In fact, let us choose on Γ a complete g_n^{n-p} and let us consider its projective image L, which will be a curve of order n belonging to the space S_{n-p} and birationally equivalent to Γ. Since L does not possess multiple points,* a generic pencil of hyperplanes cuts out on L a g_n^1 satisfying the required conditions. In fact, the ruled surface R, generated by the lines joining the pairs of points belonging to groups of the g_ν^1 (image on L of the g_ν^1 on Γ), possesses $d = (n-1)(\nu-1) - p$ distinct generators which meet a generic S_{n-p-2}, and these are neither on multiple points of the g_ν^1, nor on sets of three or more points belonging to groups of the g_ν^1. The first statement follows from the fact that the multiple points of the g_ν^1 are finite in number; the second statement is a consequence of the fact that no generator of R is on more than two points of L, since, if L possessed a trisecant, the g_{n-1}^{n-p-1}, cut out on L by the hyperplanes through one of the three meet points, would possess a neutral pair, which is impossible, since its order $n-1$ is greater than $2p$.† If, in addition, we choose the S_{n-p-2} in such a manner, that it should not meet the virtual singular tangents of L, the g_n^1 cut out on L by the pencil on S_{n-p-2} will also satisfy the condition of possessing double points only.

If we now consider on Γ the g_n^1 homologous to the g_n^1 just constructed on L, the groups of the g_n^1, g_ν^1, which pass through a generic point of Γ, do not have other points in common, and two groups of g_n^1, g_ν^1, which have more than one point in common, have in common two points only, which are distinct. Let G, G' be two (generic) groups of g_n^1, g_ν^1, which are made up of distinct points and have no points in common, and such that the n (and the ν) groups of the g_ν^1 (and of the g_n^1), which pass through the points of G (and of G' respectively), be distinct (and also, if desired, do not

* See my *Vorlesungen über algebraische Geometrie*, Leipzig, Teubner (1921), p. 134, or also my *Trattato di geometria algebrica*, Bologna, Zanichelli (1926), p. 148.

† *Ibid.*

contain multiple points). Let us set up a projectivity π between the groups of the g_n^1 and the lines of the pencil P and a projectivity π' between the groups of g_ν^1 and the lines of the pencil Q, in such a way that to the line PQ there correspond in π and in π' the groups G, G' respectively. The curve C_0, the locus of intersections of pairs of lines on P, Q, which correspond in π, π' to groups of g_n^1, g_ν^1 passing through a variable point on Γ, satisfies, as a member of the system H of the analogous curves C, the conditions 1), 2).* From 1) and 2) it follows, that *the characteristic series of K, on C_0, is complete.* In fact, let i' be the index of speciality of the two-fold of the g_ν^1 cut out on C_0 by the lines of the pencil Q. Under this hypothesis, among the points of contact M_0 of the σ lines λ, drawn through P and touching C_0, there are $\sigma - i'$—we shall call these points N_0—which impose independent conditions on the curves D passing through the double points of C_0, so that the D's which pass through the points N_0 contain all the points M_0. We denote by Σ' the system of the curves D passing through the double points of C_0. The $\sigma - i'$ tangents relative to the points N_0 will then impose $\sigma - i''$ ($i'' \geqq i'$) conditions on the C's of H constrained to touch them, so that these C's will form a system $\bar{K}$ of dimension

$$(2n + 2\nu + p - 1) - (\sigma - i'') = 2\nu - p + 1 + i''.$$

The $\infty^{2\nu-p+i''}$ curves of $\bar{K}$, which are infinitely near to C_0, are curves D of Σ' passing through the points N_0 and consequently through all the points M_0. Hence the said curves of $\bar{K}$ cut out on C_0, outside of fixed points, groups of the characteristic series of K, which series (completed, if necessary) has the dimension $2\nu - p + i'$. Hence

$$2\nu - p + i'' \leqq 2\nu - p + i', \text{ or } i'' \leqq i',$$

and since $i'' \geqq i'$, it follows that $i'' = i'$.

Reasoning as in section 5, we deduce that the curves of $\bar{K}$ (here, and also in the analogous cases considered below, we mean of course: of an irreducible part of $\bar{K}$ which contains C_0) touch the remaining i' lines λ, and consequently $\bar{K}$ coincides with the system K, which has thus the dimension $2\nu - p + 1 + i'$. It follows that the characteristic series of K on C_0, and hence also on a generic curve of K, is complete.†

* Cf. also in this connection my *Vorlesungen*, p. 100.

† Cf. the concluding remarks of this letter, which show that *any* curve satisfying the conditions 1), 2) belongs to only one irreducible system K.

The curves of K are birationally equivalent to C_0.* From this it follows that there exists on Γ an irreducible variety W of dimension $2(\nu-1)-p+i'$ of series g_ν^1, and that the two-fold of a generic of these series has the index of speciality i'. In fact, since the (complete) characteristic series of K on a generic curve C of the system has the dimension $2\nu-p+i'$ and—consequently—the index of speciality i', it follows that i' is the index of speciality of the two-fold of the g_ν^1 cut out on the above C by the lines through Q. To this g_ν^1 on C there corresponds a g_ν^1 on Γ. Thus with every C in K there is associated a g_ν^1 on Γ, and if two curves C_1, C_2 of K are associated with the same g_ν^1 on Γ, the groups of the g_ν^1's cut out on C_1, C_2 by the lines on Q correspond to each other in a projectivity, homologous groups being those, which arise from one and the same group of the g_ν^1 on Γ in the birational correspondences between Γ, C_1 and Γ, C_2. Hence the distinct g_ν^1's on Γ, associated with the curves C of K, depend on

$$(2\nu-p+1+i')-3=2(\nu-1)-p+i' \text{ parameters.}$$

Now, considering that the g_n^1's on Γ constitute an (irreducible) variety of dimension $2(n-1)-p$ and that to every pair of series g_n^1, g_ν^1, assigned on Γ (such that g_ν^1 is in W), there correspond ∞^6 curves C birationally equivalent to C_0 and *contained, as C_0 itself, in the system H'*, we arrive immediately at the formula (1).

I now make some additional remarks which lead to the more expressive theorem relative to the case $i>0$. I prove in the first place that:

Any g_ν^1 on a curve Γ of genus p, without fixed points, belongs to a unique complete (irreducible) variety of series of the same order. Consequently two such varieties cannot have in common other than series with fixed points.

Let us denote by i the index of speciality of the two-fold of the g_ν^1 (i is not less than the greater of the two numbers 0, $p-2\nu+2$), and let C_0 be, as above, the curve constructed by means of this g_ν^1 and a generic g_n^1 on Γ ($n>2p+1$). We denote by N_0 the points of contact of $\sigma-i$ tangents λ, drawn through P, with C_0, which impose independent conditions on the curves of Σ' constrained to contain them. The curves of H, which are infinitely near to C_0 and which touch the above $\sigma-i$ lines in neighboring points of N_0, will then satisfy $\sigma-i$ independent conditions and will automatically touch the remaining lines λ. Hence these curves constitute the neighborhood I of C_0 in K. The curves of H, which touch one of the above

* See my *Vorlesungen*, p. 340, where this statement is proved in an analogous case. The same reasoning holds also here.

$\sigma - i$ lines in a point of the neighborhood of the corresponding point N_0, constitute a linear analytical branch. The $\sigma - i$ linear regions of origin C_0, which are thus obtained, possesss at C_0 independent tangent spaces. This shows that the neighborhood I consists of *only one linear branch* of origin C_0, whence the theorem follows.

As a corollary we have the following theorem:

If a g_ν^1, without fixed points, on a Γ of genus p belongs to a complete (irreducible) variety W of series of the same order, and if W has the dimension $2(\nu - 1) - p + i$, then the two-fold of g_ν^1 has the index of speciality i. In fact, the index of speciality of a generic g_ν^1 is i, because the system K derived from a *fixed* generic g_n^1 $(n > 2p + 1)$ and the variable g_ν^1 in W, by fixing the projective correspondence between the groups of the g_n^1 and the lines on P, is such that its characteristic series is complete. On the other hand, if C_0 is the curve which corresponds to a particular g_ν^1 without fixed points, the system of the curves C of H, which touch the lines λ, is such that its characteristic series, on C_0, is complete. But by the previous theorem that part of this system, which contains C_0, must necessarily coincide with K, and hence the theorem follows. It should be noticed that it also follows that in (1) the sign $=$ holds.

* * * * *

The proof of uniqueness, which makes use of the method of intersection of analytical branches, given by me in the "Vorlesungen," has a wider scope. It serves namely to prove, that when a continuous system K of plane curves C, of order m, is defined by certain linear conditions (such that the curves C satisfying these conditions only constitute a linear system Σ of dimension ρ) and by the condition that the C's have d nodes and σ simple contacts with a given curve Δ (which may also be reducible), then a curve C_0 of K, which possesses d nodes and σ simple contacts with Δ and *on which moreover the characteristic series of K is complete,* is the origin of a unique linear branch of the system. In fact, to say that the characteristic series of K on C_0 is complete, is the same as to say,* that the number $\sigma - i$ $(i \geqq 0)$ of conditions, imposed on the adjoint curves of order m of C_0 by the σ points of contact M_0 of C_0 with Δ, is equal to the number of conditions imposed by σ contacts with Δ on the curves of order m, possessing d ordinary double points and sufficiently near to C_0. Moreover it is possible to determine $\sigma - i$ of the points M_0—let us call them N_0—, such that the

* Cf. my Note, "Sulla teoria degl'integrali semplici di 1ª specie appartenenti ad una superficie algebrica," *Rendiconti della R. Accademia dei Lincei,* (1921), p. 297.

curves, which are near to C_0 and which touch Δ in points near to the points N_0, should necessarily touch Δ in points which are near to the remaining i points M_0. Now, in the space S_ρ of the curves of Σ we have, that the elements near to C_0, which possess a double point in the neighborhood of a double point of C_0, and the elements near to C_0, which have a contact with Δ in the neighborhood of a point N_0, constitute as many linear branches of origin C_0 and of dimension $\rho - 1$. Thus the totality of all curves of Σ near to C_0 is obtained as the intersection of $d + \sigma - i$ linear branches of origin C_0, the tangent hyperplanes of which at C_0 are linearly independent. Hence the intersection of these branches is *one* linear branch of dimension $\rho - d - \sigma + i$.

In the case in which Δ is the branch curve of a multiply covered plane, we are led to the following theorem:

Given on an algebraic surface F two continuous systems of curves of the same order and with the same number of (variable or fixed) nodes and such that the characteristic series of the systems on their respective generic curves are complete, the only curves which the two systems can have in common are those, which possess a larger number of double points, or those, on which the characteristic series is not complete.

This theorem is found in my Note of 1916,* save the last alternative, which is omitted there, since at that time the proof of the completeness of the characteristic series of a complete continuous system (Enriques) had not yet encountered the objection which I later raised against it in the quoted Note of 1921. As you know, I have proved in the same paper, by transcendental methods, that on an algebraic surface F there always exist continuous systems of curves the characteristic series of which is complete. However, it would be of great importance to prove this theorem by algebro-geometric methods. It is necessary, for this purpose, to make a thorough study of complete continuous systems of plane curves in the definition of which also occur contact conditions. These are the systems which, for the above purpose, are of most interest and for which it is not possible to affirm, at the present stage of our knowledge, that the characteristic series is complete, *even if we deal with curves possessing ordinary double points only.* The exception to Enriques' theorem in regard to systems of curves with higher singularities is of less importance, because its bearing on the geometry of the surfaces is almost insignificant. In many cases (as you have shown in an interesting example in your paper) this exception can even be removed,

* "Nuovi contributi alla teoria dei sistemi continui, etc.," *Ibid.* (1916), p. 469.

by considering the system as the limit of a system with ordinary double points only and by properly modifying the definition of a characteristic series.

I conclude with two theorems which follow immediately from the preceding considerations:

a) *On a curve* Γ *the* g_n^r*'s* $(r \geqq 1)$ *without fixed points are distributed into a finite number of continuous systems, any two of which have in common* g_n^r*'s with fixed points.*

b) *On an algebraic surface* F, *of geometric genus* $p_g > 0$, *the generic curve of any linear irreducible system* $|C|$, *which contains partially the canonical system* $|K|$ *of* F, *is of special moduli.*

In fact, if we put $C = K + D$, we will have that $|C'| = |C + K| = |D + 2K|$ is the adjoint of $|C|$. Hence the series cut out on C by $|2K|$ is special, and consequently the generic C is of special moduli, since it contains a series $|(C, K)|$ the two-fold of which is special.

ROME, FEBRUARY 27, 1930.

Grundlagen der kombinatorischen Logik.

TEIL I.

von H. B. CURRY.

INHALTSÜBERSICHT.

KAPITEL I. ALLGEMEINE GRUNDLAGEN.

A. EINLEITUNG.

Im Anfang jeder mathematisch-logischen Untersuchung setzt man eine gewisse Menge von Kategorien voraus, die als ein Teil des irreduziblen Minimums von Kenntnis, womit man anzufangen hat, betrachtet werden sollen, Als solche Kategorien gelten gewöhnlich Aussagen und Aussagefunktionen von verschiedenen Ordnungen und Stufen. Diese Kategorien müssen weiterhin nicht nur als rein formale Begriffe vorausgesetzt werden, sondern es muss zuweilen einen Sinn haben, dass ein gegebener Gegenstand zu dieser oder jener Kategorie gehört. Mit anderen Worten, sie müssen eine inhaltliche Eigenschaft besitzen, nämlich die, dass sie Kategorien sind; und so sind sie das, was ich schon nicht-formale Grundbegriffe genannt habe.†

* Die Abschnitte *C*, *D*, *E* werden in Teil II erscheinen.

† In einer Abhandlung, "An Analysis of Logical Substitution," *American Journal of Mathematics*, Bd. 51 (Juli, 1929), S. 363-384.

Die mit diesen Kategorien verbundene vorausgesetzte Kenntnis reicht aber weit über diese Eigenschaft hinaus. In der Tat müssen wir neben den Kategorien auch gewisse Verknüpfungen annehmen—die natürlich wieder nicht-formale Grundbegriffe sind, indem sie Verfahren angeben, wonach zwei oder mehrere Gegenstände kombiniert werden können, um einen neuen Gegenstand hervorzubringen. Unter diesen Verknüpfungen gehen uns hier alle Arten von Substitutionsprozessen an. Davon muss man die folgenden Eigenschaften anfänglich erkennen: erstens diejenigen, welche uns die Möglichkeit eines Substitutionsprozesses kundgeben, wenn nur die betreffenden Gegenstände zu gewissen Kategorien gehören; zweitens die, woraus wir unter Umständen schliessen, dass zwei verschiedene Substitutionsprozesse dasselbe Resultat liefern *; drittens diejenigen, welche die Kategorie des Ergebnisses bestimmen. Diese Kategorien, Verknüpfungen und Eigenschaften bilden daher eine vorausgesetzte Lehre von beachtenswerter Verwicklung; und diese Lehre möchte ich die *Urlogik* der betreffenden Theorie nennen.

Trotz des fundamentalen Charakters dieser Urlogik für irgendeine existierende mathematische Theorie der Logik, haben doch einige der wichtigen Probleme schon in ihr ihren wesentlichen Ursprung. Ich betrachte hier die zwei folgenden:

1) Die Vereinfachung der Grundlagen. Das Wesen dieses Problems scheint mir darin zu liegen, dass man die gesamte Mathematik und Logik auf ein Minimum von Urkenntnis aufbauen will; oder genauer, dass man sie in ihre Elemente zerlegen will, damit diese Urkenntnis mit schärfster Klarheit und Deutlichkeit ausgeprägt wird. Aus diesem Grunde strebt jedermann vermutlich danach, die Anzahl der Grundbegriffe und Axiome zu vermindern; und zwar sind einige Logiker so weit in dieser Richtung gegangen, dass sie nicht davor zurückschrecken, Allgemeinheit, Vollständigkeit und sogar Exaktheit diesem Streben zu opfern.† Allein für die Logik ist die Anzahl der Grundbegriffe und Axiome dieser Art eine Betrachtung von geringerer Wichtigheit. Denn was macht es eigentlich, dass wir zwei oder drei Grundbegriffe ausschalten, wenn schon in der Urlogik selbst unendlich viele Grund-

* Vgl. Kapitel II unten; auch für einfache Beispiele den Teil II meiner oben zitierten Abhandlung.

† Z. B. müssen wir in den *Principia Mathematica*, Implikation und den bestimmten Artikel als Grundbegriffe entbehren. Die dort gegebenen Definitionen dieser Grundbegriffe führen zu den paradoxen Resultaten. In der Tat ist dort der König von Frankreich von dem König von Frankreich verschieden; und zwischen den zwei Aussagen: 1) "hätten alle Menschen drei Hände, so würde auch Bismarck drei Hände haben," 2) "hätten alle Menschen drei Hände, so wäre der Mond aus grünem Käse aufgebaut," wird kein Unterschied gemacht.

begriffe und Regeln, und zwar nicht formale, vorhanden sind? Die Vereinfachung der Grundlagen soll also mit der Vereinfachung der Urlogik beginnen.

2) Die Beseitigung der Antinomien. Um uns die Verwandtschaft von diesen mit der Urlogik zu vergegenwärtigen, betrachten wir z. B. das Russellsche Paradoxon, das in der folgenden Weise aufgefasst werden kann: F sei diejenige Eigenschaft von Eigenschaften, für welche

$$F(\phi) = \text{nicht } \phi(\phi);$$

dann ist
$$F(F) = \text{nicht } F(F).$$

Behauptet man, dass $F(F)$ eine Aussage sei, wo eine Aussage als etwas, das entweder wahr oder falsch ist, definiert wird, so hat man sofort einen Widerspruch. Aber wir können den Widerspruch allerdings dadurch vermeiden, dass wir in Abrede stellen, dass $F(F)$ zu der Kategorie der Aussagen, oder F selbst zu der der Eigenschaften gehöre. Gerade hier stossen wir auf einen Lehrsatz der Urlogik. Und zwar sind es diese Antinomien, die uns zwingen, eine entwickeltere Urlogik uns anzueignen, als sonst begreiflich wäre. Es mag also sein, dass ein tieferes Studium der Urlogik Licht in dieses dunkle Gebiet der Antinomien verbreiten wird.

Diese Umstände haben mir die Anregung gegeben, jene Urlogik mathematisch zu behandeln. D. h. genauer: eine abstrakte oder formale Theorie aufzubauen, welche auf eine sehr einfache Urlogik gegründet wird, und durch welche die Fragen, die man in der gewöhnlichen anschaulichen Urlogik zu stellen pflegt, durch symbolische Beweisführung beantwortet werden können. In dieser Abhandlung beabsichtige ich eine solche Theorie zu begründen. Weil die betreffenden Fragen einen wesentlich kombinatorischen Charakter haben, habe ich die Theorie *kombinatorische Logik* genannt.

Diese kombinatorische Logik wird wohl fähig sein als Grundlage einer abstrakten Theorie der gesamten Logik und Mathematik, einschliesslich Funktionen (Prädikaten, Relationen) von beliebig vielen Variablen, zu dienen. In der Tat bin ich der Ueberzeugung, dass eine solche Theorie durch Hinzufügen von endlich vielen formalen Grundbegriffen und Axiomen zu dem unten gegebenen Grundgerüst* gegründet werden kann. Weil in einer so eingerichteten Theorie das Grundgerüst überhaupt nur endlich viele Bestandteile hat, und weil ferner die Regeln nur ungefähr denselben Grad von Zu-

---

* Dieses Wort ist eine Uebersetzung der englischen Phrase "primitive frame," die ich in der oben erwähnten Abhandlung definiert habe. Es bedeutet die Gesamtheit der vorausgesetzten Grundbegriffe, Axiome und Regeln.

sammengesetztheit wie die wohlbekannte Schlussregel haben, so wird dabei ein wesentlicher Fortschritt gegen das erste oben erwähnte Problem erzielt. Was das zweite betrifft, kann ich zunächst darüber nichts aussagen.

Das in dieser Abhandlung ausgeführte Problem ist doch nur ein Teil der kombinatorischen Logik überhaupt. Dieser Teil ist die Analysis der Substitutionsprozesse, formell betrachtet und abgesehen von den Kategorien, wozu die Gegenstände gehören; d. h. ich will die zweite der drei im zweiten Paragraphen erwähnten Arten von Eigenschaften untersuchen.*

Der Weg zu dieser Analysis ist von M. Schönfinkel † angedeutet worden. Er versteht zunächst eine Funktion in einem etwas neuen Sinne, nämlich als eine Zuordnung eines Funktionswerts zu jedem Element eines Argumentbereiches, wo sowohl der Funktionswert, als auch das Argument eine Funktion sein kann. Er will dann eine Funktion von n Variablen (im gewöhnlichen Sinne) als eine Funktion (im neuen Sinne) betrachten, deren Funktionswert für das Argument x eine Funktion (im gewöhnlichen Sinne) von $(n-1)$ Variablen ist, und zwar diejenige, welche durch Einsetzung von x in die erste ‡ Leerstelle der ursprünglichen Funktion entsteht. Mit dieser Auffassung verknüpft er die folgende Bezeichnungsweise. Leerstellen gebraucht er nicht, sondern er bezeichnet Funktionen wie andere Dinge mit einfachen Buchstaben (oder Gruppen von Buchstaben). Den Funktionswert einer Funktion für ein gegebenes Argument bezeichnet er weiterhin damit, dass er das Zeichen für das Argument gleich rechts von dem für die Funktion geltenden schreibt. Zur Erläuterung dieser Ideen schreibe ich gleich unten links, einige Ableitungen von einer Funktion f, die wir im gewöhnlichen Sinne als eine Funktion von drei Variablen begreifen, und rechts dieselben Gegenstände in etwas anschaulicherer Schreibweise:

$$\begin{array}{ll} f & f(-_1,\ -_2,\ -_3) \\ fx & f(x,\ -_1,\ -_2) \\ (fx)y & f(x,\ y,\ -_1) \\ ((fx)y)z & f(x,\ y,\ z) \end{array}$$

Schönfinkel setzt weiter fest, dass in einem Ausdruck wie $\cdots(((fx_1)x_2)x_3)\cdots$ die die letzten Seiten mit umfassenden Klammern wegbleiben dürfen.

* Die erste Art wird auch erledigt. In der Tat sind diese Eigenschaften nach den Betrachtungen von B 1 (unten) ganz trivial.

† *Mathematische Annalen*, Bd. 92 (1924), S. 305-316.

‡ Es ist allerdings hier vorausgesetzt, dass die Leerstellen in einer durch die Funktion selbst bestimmten Weise numeriert sind. Dieselbe Annahme liegt doch der Schönfinkelschen Darstellung zugrunde. Vgl. unten II, A.

Um nun etwas allgemeinere Ausdrücke darzustellen, setzt er einige bestimmte Funktionen fest. Diese möchte ich hier mit den Buchstaben I, K, B, C, S bezeichnen.* Sie sind dadurch definiert, dass für beliebige Gegenstände x, y, z die folgenden Definitionsregeln gelten:

$$
\begin{aligned}
Ix &= x \\
Kxy &= x \\
Bxyz &= x(yz) \qquad (1)\\
Cxyz &= xzy \\
Sxyz &= xz(yz).
\end{aligned}
$$

Vermöge dieser Funktionen ist er in der Lage, andere Ableitungen einer Funktion von drei Variabeln ohne Leerstelle darzustellen, etwa

$$
\begin{array}{cc}
Cfx & f(-_1,\ x,\ -_2) \\
C(BCf)xy & f(y,\ -_1,\ x) \\
Bfg & f(g(-_1))
\end{array}
$$

u. s. w. In der Tat gilt nun der allgemeine Satz: Wenn irgendein logischer Ausdruck vorhanden ist, der aus gewissen Konstanten $u_1, u_2, \cdots u_m$ und numerierten freien Variablen zusammengesetzt ist, so kann er in der Schönfinkelschen Schreibweise in der Form

$$(2) \qquad Yu_1, u_2 \cdots u_m$$

wo Y eine Zusammensetzung von nur I, K, B, C, S ist, dargestellt werden. Diese Darstellung bedeutet, dass, wenn man nach (2) die Variablen in ihrer natürlichen Reihenfolge hinzufügt, und dann die Funktionen I, K u. s. w. durch die Definitionsregeln (1) eliminiert, der resultierende Ausdrück immer mit dem zuerst gegebenen identisch ist.—Dieser Satz ist von Schönfinkel nicht bewiesen, aber er enthält vom hier vorliegenden Gesichtspunkte aus das Wesen seiner Gedanken. Der Satz ist hier allerdings etwas roh dargelegt; genauere Sätze werden aber unten im formalen Teile streng bewiesen.

Nun macht Schönfinkel Anspruch darauf, dass er die Begriffe Variable, Aussage und Aussagefunktion aus der Logik entfernt habe. In dieser Hinsicht sind aber folgende Betrachtungen zu bemerken. Erstens: das Y des eben dargelegten Satzes ist doch nicht eindeutig bestimmt; z. B. sind $BCCf$ und If in dem Sinne identisch, dass sie denselben Ausdruck darstellen. Aber diese Identität ist keineswegs aus den Definitionsregeln zu beweisen; zu diesem Behuf muss man vielmehr die Variablen tatsächlich hinzufügen. Infolge-

* **Schönfinkel hat diese bzw. mit I, C, Z, T, S bezeichnet.**

dessen sind doch die Variablen, obgleich sie in den Formeln selbst nicht erscheinen, mit allen ihren assoziierten Grundbegriffen noch unvermeidbar. Zweitens: Die Begriffe Aussage und Aussagefunktion sind auch nur scheinbar eliminiert. Denn jede Funktion hat ihren Argumentbereich, und wie definiert man diesen, wenn man jene Begriffe nicht versteht? Die allgemeinen Sätze der Logik gelten ja gewöhnlich nur, wenn die Variablen zu bestimmten Kategorien gehören. Z. B. ist der Satz von der Identität $A \rightarrow A$, nur wahr, wenn A eine Aussage ist. Weil Schönfinkel keineswegs gezeigt hat, wie die Einführung jener Grundbegriffe zu vermeiden ist, und weil er sie nicht aus anderen definieren kann, so hat er seinen Anspruch nicht gerechtfertigt. Er hat in der Tat nur eine neue und unbequemere Schreibweise gewonnen.

Jedoch führen diese Gedanken Schönfinkels zur Lösung des oben vorgelegten Problems. Denn die Funktionen B, C, u. s. w. bilden die Elemente, in die die Substitutionsprozesse sich zerlegen lassen. (Ich werde nachher genau zeigen, wie der allgemeinste solche Prozess sich aus diesen Funktionen zusammensetzen lässt). Das Wesen des Problems liegt daher in den im ersten Teile des letzten Absatzes besprochenen Identitäten. Es lässt sich wohl auf die folgende Aufgabe zurückführen: das Grundgerüst so festzustellen, dass jede dieser Identitäten rein formal bewiesen werden kann. Diese Aufgabe wird unten vollständig gelöst,* und zwar so, dass 1) die Funktionen $B, C, \cdots$ nur als rein formale Begriffe auftreten, und 2) die Variablen in den formalen Beweisen nicht zu erscheinen brauchen. Der Beweis, dass jede solche Identität aus dem Grundgerüst ableitbar ist, macht das Hauptergebnis des zweiten Kapitels aus. Dort wird auch auseinandergesetzt, wie dies die Lösung des früheren Problems liefert.

Diese Untersuchung habe ich immer in Rücksicht auf eine allgemeine Theorie der Logik durchgeführt. Dieses Ergebnis betrachte ich nur als eine Vorbereitung zu einer weiteren Theorie, die ich später fortzuführen hoffe.

B. Einige Philosophische Betrachtungen.

Bevor ich zum Aufbau der formalen Theorie übergehe, möchte ich hier einige philosophische Vorbemerkungen darlegen, die für die Theorie massgebend gewesen sind. Diese betreffen die logische Auffassung, welche der Theorie vorangeht und durch die Ausdeutung der letzteren heranreift. Mit der formalen Theorie als solcher haben sie natürlich nichts zu tun; und der Leser, der sich nicht dafür interessiert, kann sie ruhig überspringen.

* Einige Anfänge dazu befinden sich in meiner früheren Abhandlung. Die vorliegende Behandlung ist allgemeiner und in sich vollständig; ich weise auf die frühere nur für gewisse Einzelheiten hin.

§ 1. *Sinnlose Begriffe.* In der gewöhnlichen Logik kommen gewisse Gegenstände (Begriffe, Dinge) vor, die man als sinnlos zu titulieren pflegt. Ich möchte nun gerade fragen, was dies bedeutet. Man darf wohl behaupten, dass ein Wort oder ein Zeichen in bezug auf eine Sprache noch nicht definiert ist. Z. B. ist das Wort "cow" auf deutsch sinnlos, doch gelegentlich anf englisch sinnvoll. Aber in den vorhandenen Fällen sind es nicht Worte, sondern Begriffe, die sinnlos sein sollen. Z. B. behaupten Whitehead und Russell (und auch andere), dass $\phi(\phi)$ für jedes ϕ bedeutungslos sei. Zu sagen, dass diese Behauptung nur das Zeichen $\phi(\phi)$ betrifft, heisst, die Frage zu umgehen; weil es doch erstens ein gedachtes Etwas gibt, das das Zeichen $\phi(\phi)$ den Konventionen gemäss bedeuten mag, und zweitens, weil die Gründe,—nämlich das "Vicious Circle Principle" u. s. w. —, woraus die Behauptung entsteht, mit dem Zeichen gar nichts zu tun haben. Die Sinnlosigkeit scheint doch einen sachlichen Inhalt zu haben, aber welchen, lässt sich nicht sagen.

Zunächst gibt es aber einen Sinn, worin alles Denkbare eine Bedeutung hat, nämlich als Begriff. Hier ist ein Begriff als irgendetwas zu verstehen, das identifiziert und von anderen Dingen unterschieden zu werden vermag.* Dann ist es allerdings unbedingt Unsinn zu sagen, dass etwas nicht als Begriff existiere; denn bevor man einen solchen Satz verstehen konnte, muss man sich die Sache als Begriff schon vorgestellt haben. Sogar die "sinnlosen" Gegenstände sind also Begriffe und haben als solche eine Bedeutung.

Unter den Begriffen kommen aber einige vor, die "in sich widerspruchsvoll" sind. Solche sind das oben erwähnte $F(F)$, die grösste Kardinalzahl, die kleinste undefinierbare Ordinalzahl, das kreisförmige Quadrat u. s. w. Diese Begriffe sollen wegen der Widersprüche bedeutungslos sein. Aber die Widersprüche liegen nicht in den Begriffen selbst, sondern in den Eigenschaften, die man ihnen zuordnen möchte: z. B. führt das oben erwähnte $F(F)$ zu einem Widerspruch nur dann, wenn man behauptet, dass es eine Aussage sei; die grösste Kardinalzahl nur dadurch, dass sie wirklich eine Kardinalzahl sei u. s. w.

Betrachten wir nun die sinnlosen Gegenstände überhaupt. Bezieht sich nicht derselbe Gedanke auch auf sie? Ja, die Sinnlosigkeit dieser Begriffe besteht nur darin, dass es Eigenschaften gibt, die sie nicht besitzen. Und zwar, dürfte ich genauer sagen, dass sie den gewöhnlichen Kategorien nicht

* Es ist wohl bemerkenswert, dass ein Begriff nach dieser Definition ein Gegenstand und nicht ein Prozess des Denkens ist. Beispiele von Begriffen sind etwa Bismarck, Göttingen, Tier, Regenschirm, rot, Temperatur, Materie, Substanz, Kausalität, Etwas, Funktion, der König von Frankreich, die grösste Kardinalzahl u. s. w.

angehören. Diese Kategorien sind in der Tat als etwa inhaltliche Grundbegriffe vorausgesetzt, und nichts wird betrachtet ausser was dazu gehört. Natürlich müssen Begriffe, die mit dem Wesen dieser Kategorien unverträgliche Eigenschaften besitzen, als "sinnlos" aus der Theorie ausgeschlossen werden. Aber gerade in diesem Ausschliessen besteht ein Mangel. Die Aufgabe der Logik ist die Erklärung des Denkens; wenn es von der Erklärung ausgeschlossenes Denken gibt, so ist sie fehlerhaft. Weiterhin sind es genau diese sinnlosen Begriffe, die zu Widersprüchen führen; wenn man sie ausschliesst, so kann man wohl die Antinomien vermeiden, aber nie erklären. Dass etwas ein Begriff ist, ist das einzige Erfordernis, damit man das Ding in der Logik behandeln könne.

Die Kategorie Begriff—oder, wie ich sie nachher nennen werde, um gewisse Nebenbedeutungen zu vermeiden, Etwas,—ist daher die grundlegende Kategorie der Logik überhaupt. Diese Kategorie ist ein Begriff, und seine blosse Betrachtung ist doch in sich widerspruchsfrei. Aus diesen Gründen habe ich sie als Grundkategorie der Theorie vorausgesetzt. Daraus folgt eine wichtige Konsequenz: ich brauche mich nicht mehr, mindestens insofern als es die Einführung neuer Gegenstände in die Theorie betrifft, um die Definitionsbereiche von Funktionen zu kümmern. Wenn ich z. B. den Schönfinkelschen Funktionsbegriff als Grundverknüpfung — und Schöfinkel hat schon gezeigt, dass dies allein notwendig ist—, voraussetze, so ist die Zusammensetzung von irgend zwei Begriffen, etwa f und x zu fx, wieder ein Begriff. Z. B. ist die Zusammensetzung des Königs von Frankreich mit der Aussage "der Mond ist aus grünem Käse aufgebaut," ein Begriff, weil sie identifiziert und von anderen Dingen unterschieden zu werden vermag. Unter den so hergestellten Begriffen werden einige "sinnvoll," andere "sinnlos" sein; die Hauptaufgabe der kombinatorischen Logik ist wohl, diese zwei Arten zu unterscheiden.*

§ 2. *Der Vorrang des Aussagenkalküls.* In den heutigen logischen Theorien bildet der sog. Aussagenkalkül den Grundbestandteil. Ich möchte hier die vielleicht triviale Bemerkung darlegen, dass dieser Vorgang nicht notwendig ist. Natürlich beginnt die Logik mit Sätzen, die wir fähig sein müssen, zu verstehen. Aber daraus folgt ebensowenig, dass man mit der allgemeinen Theorie von Aussagen anfangen muss, als aus der Tatsache, dass wir es im Aussagenkalkül mit gewissen Aussagefunktionen zu tun haben, folgt, dass wir mit Aussagefunktionen beginnen müssen. Weiterhin sind

* Dabei wird die erste Art von Eigenschaften die ich in I A erwähnt habe, auf die Dritte zurückgeführt.

die Begriffe "Aussage" und "Behauptung" verschieden; dieser ist mit der Wahrheit eng verbunden, jener eine Kategorie von bloss betrachteten Etwasen. Was wir im Anfang verstehen müssen, ist erstens, dass die Sätze Behauptungen sind, und zweitens, dass sie sich vermöge der Regeln zu neuen Behauptungen umgestalten lassen. Was wir mit solchen einfachen Anfängen machen können, zeigt die hier gegebene Theorie an.

Der Begriff von Behauptung muss natürlich als Grundbegriff angenommen werden und zwar als ein solcher, der in der formalen Theorie einem nicht formalen Begriffe entspricht. Diesen nicht formalen Begriff habe ich "*Formel*" genannt.

§ 3. *Unbeschränkte Universalen.* In Nummer 1 habe ich behauptet, dass der Begriff "Etwas" als grundlegender Begriff angenommen werden darf. Ich behaupte nun ferner, dass es Eigenschaften gibt, die für jedes Etwas gelten. Gegen die Möglichkeit solcher Behauptungen steht aber das wohlbekannte Verbot der nicht prädikativen Begriffsbildungen. Dieses Verbot hat im allgemeinen einen etwa pragmatischen Charakter, und lässt sich folglich nur dadurch widerlegen, dass man wirklich eine dem Verbot widersprechende Theorie aufbauen kann. In den "Principia Mathematica" wird aber dieses Prinzip in der Form des "Vicious Circle Principle," a priori verteidigt. Das dortige Argument ist etwa dies: man könnte nichts über alle Aussagen behaupten, weil dadurch neue Aussagen erschaffen würden, und also es keine bestimmte Gesamtheit von Aussagen gäbe, die den Wirkungsbereich der behaupteten Eigenschaft bildete. Ebensogut aber ist dieses Argument: man könnte nichts über alle Apfelsinen behaupten, weil jedes Jahr neue Apfelsinen erschaffen werden, und es also keine bestimmte Gesamtheit von Apfelsinen gäbe u. s. w. u. s. w. In der Tat begreifen wir allgemeine Urteile nicht in extenso, sondern in intenso—es gibt etwas im Wesen der betreffenden Eigenschaft, wovon wir schliessen können, dass sie für jedes Etwas gilt. Dieser Schluss setzt doch keine Gesamtheit voraus. Es fehlt daher diesem Einwand gegen die Möglichkeit unbeschränkter Universalen ein theoretischer Grund. Infolgedessen mache ich die einfachere und naturgemässere Voraussetzung, dass es solche Universalen gibt; insbesondere, dass solche Eigenschaften wie das reflexive Gesetz der Gleichheit für jedes Etwas gelten.

Wenn wir aber diese Annahme machen, so werden die Grundregeln der Theorie erheblich einfacher. Denn diese Regeln sind Universalen; wenn wir sie als unbeschränkte Universalen auffassen können, so brauchen wir keine Unterscheidungen zwischen verschiedenen Arten von Begriffen inbezug auf die Regeln zu machen. Eine ähnliche Vereinfachung betrifft das ganze Grundgerüst.

§ 4. *Theorie und Metatheorie.* Die letzte Betrachtung hat die Methode der Untersuchung zum Gegenstand. Weil die im Anfang vorausgesetzten Regeln sehr einfach sind, ist eine sehr lange Darlegung erforderlich, um diese Entwicklung ausführlich zu vervollständigen. Wir interessieren uns aber nur für die Möglichkeit einer solchen Darlegung. Ich gebe demgemäss, statt die Formeln eine nach der anderen auszuschreiben, hier eine Reihe von Sätzen an, worin ich zeige, dass bestimmte Arten von Etwasen immer Formeln sind. Diese Sätze gehören eigentlich zu der Metamathematik im Hilbertschen Sinne, die mit dieser Mathematik verbunden ist. In den Beweisen dieser Sätze benutze ich nur die einfachen direkten Beweismethoden, und zwar werden die Beweise so aufgestellt, dass sie selbst, natürlich mit Benutzung der Nachweisungen, einen Prozess ergeben, wodurch der Beweis irgendeines in sie eingeschlossenen Falles ausführlich Formel nach Formel auseinandergesetzt werden kann. Unter die Methoden kommen allerdings die vollständige Induktion und andere Eigenschaften der ganzen Zahlen vor, aber dies bedeutet nicht mehr, als dass wir die ganzen Zahlen als Zeichen benutzen, und dass für einige Etwasen, die mit höheren Zahlen bezeichnet werden, jene Beweisprozesse wiederholt werden sollen. Für alle besonderen Fälle hat man eine endliche direkte Schlussfolge.

C. DAS GRUNDGERÜST.

Ich gehe jetzt zur formalen Theorie über. Die erste Aufgabe ist natürlich die Darlegung des Grundgerüsts und die Erklärung der in der Ausführung benutzten Festsetzungen.

§ 1. *Vorbereitende Erklärungen.* Die vorliegende Theorie wird vom abstrakten Gesichtspunkte aus durchgeführt; d. h. als eine Lehre über abstrakte Begriffe, die durch das Grundgerüst selbst definiert werden.* Es ist natürlich wichtig darin die formalen Entwicklungen selbst und die inhaltlichen Ueberlegungen darüber zu unterscheiden; aber ausserdem habe ich noch die Unterscheidung zwischen der Theorie selbst und ihrer Symbolik festgehalten. Jedoch kann der Leser, der einen anderen Gesichtspunkt vorzieht, wohl die Theorie durch kleine Aenderungen nach seinem eigenen Gesichtspunkt umgestalten. Z. B. ist das, was ich oben über die Endlichkeit der Anzahl von Axiomen ausgesagt habe, mit dem Gesichtspunkt verbunden, dass eine Definition nichts wesentlich Neues in die Theorie hereinbringt, und also von

* Ueber das Wesen einer abstrakten Theorie, vgl. meine oben zitierte Abhandlung.

einem Axiom verschieden ist; aber sonst kann man immer einen Satz desselben Inhalts mit etwas anderen Worten aussprechen.

In den Ausführungen brauche ich zwei Arten von Zeichen: erstens die *formalen Zeichen,* welche die Bestandteile der formalen Theorie selbst bedeuten; zweitens die *inhaltlichen Zeichen,* welche eine Bedeutung inbezug auf die inhaltlichen Überlegungen haben. Mit diesen zwei Zeichenarten sind zwei Arten von Definitionen zu betrachten; um diese zu unterscheiden, beschränke ich nunmehr das Wort *Definition* auf die Erklärungen der formalen Zeichen, während die Erklärungen der inhaltlichen Zeichen (bzw. Wörter, die in einem technischen Sinn gebraucht werden), *Festsetzungen* heissen sollen. Die Definitionen und Festsetzungen, die für die ganze Abhandlung beibehalten werden sollen, werden hervorgehoben und als solche ausgezeichnet; neben ihnen mache ich in den einzelnen Beweisen und Erläuterungen Gebrauch von solchen, die nur für den betreffenden Zusammenhang gelten sollen.

Die Definitionen werden weiterhin nun folgendermassen festgestellt. Erstens werden die Zeichen, die in der Darstellung des Grundgerüsts selbst mit gewissen Bestandteilen verbunden sind, dabei definiert. Jede spätere Definition ist dann eine Bestimmung, dass ein nicht schon in demselben Kontext definiertes Zeichen dasselbe wie ein anderes schon definiertes bedeuten soll. Die Identitätsrelation, die in diesem "dasselbe" schon vorliegt, kann, wenn man will, formalisiert werden. Definitionen dürfen sich natürlich nicht nur auf Zeichen im engeren Sinne, sondern auch auf Zeichenverknüpfungen beziehen (vgl. unten Def. 1 und 2).

In der folgenden Darlegung des Grundgerüsts gebe ich neben den zur abstrakten Theorie selbst erforderlichen Eigenschaften auch Andeutungen über die logische Interpretation des betreffenden Begriffs, Regel u. s. w. an. Diese Andeutungen sind aber in Klammern gesetzt.

§ 2. *Die nicht-formalen Grundbegriffe.* Die folgenden Begriffe seien vorausgesetzt:

a. *Etwas,* eine Kategorie (Ausdeutung, der oben in B 1 diskutierte Begriff Etwas). Etwase werden durch Buchstaben oder ähnliche Zeichen bezeichnet.

b. *Formel,* eine Kategorie (Ausdeutung: Behauptung).

c. *Anwendung,* eine dyadische Verknüpfung, d. h. eine Zuordnung, wodurch zu jedem geordneten Paar von Etwasen ein eindeutig bestimmtes Drittes zugeordnet wird. Wenn die zwei Etwase durch X bzw. Y bezeichnet sind, dann wird das zugeordnete Dritte die

Anwendung von X auf Y genannt und durch (XY) bezeichnet. (Ausdeutung: der Schönfinkelsche Funktionsbegriff, erweitert wie in B 1 oben).

§ 3. *Die formalen Grundbegriffe.* Von diesen Begriffen wird nichts vorausgesetzt, ausser dass sie Etwase sind. Ich schreibe also hier nur ihre Zeichen und Ausdeutungen:

B (Die Schönfinkelsche Zusammensetzungsfunktion).
C (Die Schönfinkelsche Vertauschungsfunktion).
W (eine Funktion derselben Art wie die vorhergehende, durch die Regel W [s. unten] definiert. Wir dürften sie die Verdoppelungsfunktion nennen).
K (Die Schönfinkelsche Konstanzfunktion).
Q (Gleichheit, d. h. logische, nicht, symbolische Identität).
Π (Allzeichen).
P (Implikation).
Λ (Konjunktion—das Und—Vgl. unten, Reg. Λ).

Die vorhergehenden Begriffe sind die einzigen, die in dieser Abhandlung erwähnt sind; für die weitere Durchführung der Logik brauchen wir natürlich gewisse andere, z. B. Aussage, Negation, Funktion, Seinzeichen (wenn nicht definiert) u. s. w.

§ 4. *Symbolische Festsetzungen.*

Festsetzung 1. Wenn X ein Etwas ist, so bezeichne ich mit $\vdash X$ den Satz, dass X eine Formel ist, und zwar sowohl wenn dieser Satz behauptet, festgesetzt oder bloss betrachtet ist. Weiterhin dürfen dann die äusseren Klammern von X, wenn es solche gibt, wegbleiben.

Festsetzung 2. Wenn X und Y Etwase sind, so bezeichne ich mit $X \equiv Y$ den Satz, dass X und Y dasselbe Etwas sind, d. h. dass X und Y als Zeichen betrachtet, dasselbe Etwas bedeuten. Unter diesen Umständen werde ich auch manchmal sagen, dass X und Y *identisch* sind. Die äusseren Klammern von X und Y dürfen auch hier wegbleiben.

Def. 1. Immer wenn $X_1, X_2, \cdots, X_n$ Etwase sind, gilt

$$(X_1X_2X_3\cdots X_n)\equiv(\cdots(((X_1X_2)X_3)\cdots)X_n).$$

Def. 2. Immer wenn X und Y Etwase sind, gilt

$$(X=Y)\equiv(QXY).$$

Links dürfen weiter die äusseren Klammern von X und Y wegbleiben, solange nur X und Y nicht selbst von der Form $(U = V)$ sind.

Def. 3. $I \equiv (WK)$.

Festsetzung 3. Die Relation zwischen zwei Etwasen X und Y, welche durch $\vdash X = Y$ ausgedrückt wird, heisst *Gleichheit*, und unter diesen Umständen sage ich, dass X und Y gleich sind.

§ 5. *Axiome.* Die hier gegebenen Axiome reichen nur für diese Abhandlung aus. Die kombinatorischen Axiome werden unten (s. II B 5)* in viel übersichtlicherer Form wiedergegeben.

a. *Axiom der Identität.*

Ax. Q. $\vdash \Pi(W(CQ))$ (Reflexives Gesetz, s. unten, Abschn. D).

b. *Kombinatorische Axiome.*

Ax. B. $\vdash C(BB(BBB))B = B(BB)B$.
Ax. C. $\vdash C(BB(BBB))C = B(BC)(BBB)$.
Ax. W. $\vdash C(BBB)W = B(BW)(BBB)$.
Ax. K. $\vdash C(BBB)K = B(BK)I$.
Ax. I_1. $\vdash CBI = B(BI)I$.†
Ax. (BC). $\vdash BBC = B(B(BC)C)(BB)$.
Ax. (BW). $\vdash BBW = B(B(B(B(BW)W)(BC))B(BB))B$.
Ax. (BK). $\vdash BBK = BKK$.
Ax. $(CC)_1$. $\vdash BCC = B(BI)$.
Ax. $(CC)_2$. $\vdash B(B(BC)C)(BC) = B(BC(BC))C$.
Ax. (CW). $\vdash BCW = B(B(BW)C)(BC)$.
Ax. (CK). $\vdash BCK = BK$.
Ax. (WC). $\vdash BWC = W$.
Ax. (WW). $\vdash BWW = BW(BW)$.
Ax. (WK). $\vdash BWK = BI$.
Ax. I_2. $\vdash BI = I$.

§ 6. *Regeln.* Die folgenden Regeln sind vermutlich für die gesamte Mathematik und Logik hinreichend. Die zwei letzteren werden nur in einem späteren Stadium benutzt, aber sie sind hier der Vollständigkeit halber gegeben.

---

* s. d. für Bemerkungen über die Bedeutung dieser Axiome.

† Dieses Axiom kann aus den anderen bewiesen werden. s. unten. II D 2, Satz 6.

Reg. E. Wenn X und Y Etwase sind, dann ist immer (XY) ein Etwas. (Ueber die Ausdeutung dieser Regel vgl. oben, B 1. Sie ist so einfach, dass sie nie wieder erwähnt wird, sondern nur implizit auftritt).

Reg. Q_1. Wenn X und Y Etwase sind, und 1) $\vdash X$, 2) $\vdash QXY$ (bzw. $\vdash X = Y$), dann $\vdash Y$.

Reg. Q_2. Wenn X, Y, Z Etwase sind, und $\vdash QXY$ (bzw. $\vdash X = Y$), dann

$$\vdash Q(ZX)(ZY) \qquad (\text{d. h. } \vdash ZX = ZY).$$

Reg. Π. Wenn X und Y Etwase sind, und $\vdash \Pi X$, dann $\vdash XY$. (Die Ausdeutung dieser Regel ist das sog. Prinzip von Aristoteles).

Reg. B. Wenn X, Y, Z Etwase sind, dann $\vdash BXYZ = X(YZ)$.

Reg. C. Wenn X, Y, Z Etwase sind, dann $\vdash CXYZ = XZY$.

Reg. W. Wenn X, Y Etwase sind, dann $\vdash WXY = XYY$.

Reg. K. Wenn X, Y Etwase sind, dann $\vdash KXY = X$.

Reg. P. Wenn X, Y Etwase sind, sodass 1) $\vdash X$, 2) $\vdash PXY$, dann $\vdash Y$. (Dies ist die wohlbekannte Schlussregel).

Reg. Λ. Wenn X, Y Etwase sind, sodass 1) $\vdash X$, 2) $\vdash Y$, dann $\vdash \Lambda XY$. (Diese Regel kann natürlich vermieden werden, wenn in den formalen Entwicklungen der Theorie geschlossen werden kann, dass für beliebige X, Y

$$\vdash PX(PY(\Lambda XY),$$

d. h. in gewöhnlicherer Schreibweise,

$$X \rightarrow (Y \rightarrow (X \& Y)).$$

Aber diese Aussage ist mit der logischen Bedeutung der Implikation unverträglich. Wenn wir diese logische Implikation eintreten lassen wollen, so müssen wir P und Λ als unabhängige Grundbegriffe, und die beiden Regeln P und Λ voraussetzen. [Vgl. Lewis, C. J.—*A Survey of Symbolic Logic.* Berkeley 1918, Chap. V]).

D. DIE EIGENSCHAFTEN DER GLEICHHEIT.

SATZ 1. *Wenn X ein Etwas ist, dann* $\vdash QXX$.

Beweis:	$\vdash \Pi(W(CQ))$	(Ax. Q).
$\therefore$	$\vdash W(CQ)X$	(Reg. Π, Hp.).
Aber	$\vdash Q(W(CQ)X)(CQXX)$	(Reg. W).
$\therefore$	$\vdash CQXX$	(Reg. Q_1).
Aber	$\vdash Q(CQXX)(QXX)$	(Reg. C).
$\therefore$	$\vdash QXX$ w. z. b. w.	(Reg. Q_1).

SATZ 2. *Wenn X und Y Etwase sind, und* $\vdash QXY$,
dann $\vdash QYX$

Beweis:	$\vdash QXY$	(Hp.)
$\therefore$	$\vdash Q(CQXX)(CQXX)$	(Reg. Q_2, $Z \equiv CQX$).
Aber	$\vdash CQXX$	(vorvorletzte Formel im Beweis von Satz 1).
$\therefore$	$\vdash CQXY$.	(Reg. Q).
$\therefore$	$\vdash QYX$. w. z. b. w.	(Reg. C und Q_1).

SATZ 3. *Wenn X, Y, Z Etwase sind, und* 1) $\vdash QXY$, 2) $\vdash QYZ$, *dann* $\vdash QXZ$.

Beweis:	$\vdash QYZ$	(Hp. 2).
	$\vdash Q(QXY)(QXZ)$	(Reg. Q_2).
$\therefore$	$\vdash QXZ$ w. z. b. w.	(Hp. 1 und Reg. Q_1).

SATZ 4. *Wenn X, Y, Z Etwase sind, und* $\vdash QXY$, *dann* $\vdash Q(XZ)(YZ)$.

Beweis:

$\vdash Q(C(B(WK))ZX)(C(B(WK))ZY)$	(Hp. und Reg. Q_2).
$\vdash Q(C(B(WK))ZX)(B(WK)XZ)$	(Reg. C).
$\vdash Q(B(WK)XZ)(WK(XZ))$	(Reg. B).
$\vdash Q(WK(XZ))(K(XZ)(XZ))$	(Reg. W).
$\vdash Q(K(XZ)(XZ))(XZ)$	(Reg. K).

aus den letzten vier Formeln und Satz 3, $\vdash Q(C(B(WK))ZX)(XZ)$. In ähnlicher Weise folgt dass $\vdash Q(C(B(WK)ZY)(YZ)$.

Aus der ersten und den zwei letzten Formeln und den Sätzen 2 und 3.
$\vdash Q(XZ)(YZ)$. w. z. b. w.

SATZ 5. *Wenn X und Y Etwase sind, und $X \equiv Y$, dann* $\vdash QXY$.

Beweis: Aus Satz 1 $\vdash QXX$.
Weil Y dasselbe wie X bedeutet, so folgt $\vdash QXY$. w. z. b. w.

SATZ 6. *Gleichheit ist eine reflexive, symmetrische und transitive Relation, und zwar derart, dass aus $\vdash X = Y$ für ein beliebiges Etwas Z folgt*

$$\vdash ZX = ZY, \text{ und } \vdash XZ = YZ,$$

und weiterhin aus $\vdash X$ und $\vdash X = Y$, folgt $\vdash Y$.

Beweis: Dieser Satz ist nur eine Zusammenfassung von Reg. Q_1 und Q_2 und Sätzen 1-4 inbezug auf C 4. Def. 2 und der Bedeutung von $\equiv$. Er lässt

sich aber auch, abgesehen von der Bedeutung von $\equiv$, nur mit Benutzung von Satz 5 ableiten. Z. B. beweise ich die letzte Behauptung.

Angenommen	$\vdash X = Y$,	
dann	$\vdash QXY$	(C 4. Def. 2, Satz 5, Reg. Q_1),
daher aus	$\vdash X$,	
folgt	$\vdash Y$	(Reg. Q_1).

SATZ 7. *$\mathfrak{X}$ sei ein Etwas, das aus gegebenen Etwasen $X_1, X_2, \cdots, X_n$ (unter anderen) durch Anwendung entsteht, und $\mathfrak{Y}$ sei der Ausdruck, welcher dann entsteht, wenn man $X_1, X_2 \cdots X_n$ in $\mathfrak{X}$ durch Etwase $Y_1, Y_2 \cdots Y_n$ ersetzt. Ferner sei für alle $i = 1, 2 \cdots n$ entweder $\vdash X_i = Y_i$ oder $\vdash Y_i = X_i$. Dann $\vdash \mathfrak{X} = \mathfrak{Y}$.*

Beweis: Wir können annehmen, dass X_i nur einmal in $\mathfrak{X}$ vorkommt, weil der Fall, dass einige X_i nochmals vorkommen, sich auf den Fall, dass sie alle nur einmal vorkommen, durch Vergrösserung des n zurückführen lässt (die X_i brauchen nicht alle verschieden zu sein). Weiterhin können wir uns auf den Fall $n = 1$ beschränken, weil es sonst immer eine Reihe von Ausdrücken $\mathfrak{X}_0 \equiv \mathfrak{X}, \mathfrak{X}_1, \mathfrak{X}_2, \cdots \mathfrak{X}_n \equiv \mathfrak{Y}$ derart gibt, dass $\mathfrak{X}_{i+1}$ sich aus $\mathfrak{X}_i$ durch Einsetzung von Y_i statt X_i erzeugt, und wenn wir dann bewiesen haben, dass

	$\vdash \mathfrak{X}_i = \mathfrak{X}_{i+1}$,	
so folgt	$\vdash \mathfrak{X}_0 = \mathfrak{X}_n$	(Satz 6),
also	$\vdash \mathfrak{X} = \mathfrak{Y}$	(Satz 5, 6).

Infolgedessen nehmen wir an, dass $\mathfrak{Y}$ sich aus $\mathfrak{X}$ durch Einsetzung von Y statt X erzeugt. Dann gibt es zwei Reihen von Etwasen $X_1, X_2 \cdots X_m$, $Y_1, Y_2, \cdots Y_m$ derart, dass

$$X_1 \equiv X, \qquad Y_1 \equiv Y,$$
$$X_{i+1} \equiv X_i Z_i, \text{ oder } X_{i+1} \equiv Z_i X_i,$$

und

$$Y_{i+1} \equiv Y_i Z_i, \text{ bzw. } Y_{i+1} \equiv Z_i Y_i,$$
$$X_m \equiv \mathfrak{X}, \qquad Y_m \equiv \mathfrak{Y}.$$

Dann haben wir zuerst $\vdash X_1 = Y_1$ (Hp., Satz 5,—vielleicht Satz 3), und weiter aus $\vdash X_i = Y_i$ folgt $\vdash X_{i+1} = Y_{i+1}$ (Satz 6).
Daher durch Induktion $\vdash X_m = Y_m$,
daher $\vdash \mathfrak{X} = \mathfrak{Y}$. w. z. b. w.

SATZ 8. *Die zwei Sätze 1) $\mathfrak{X} \equiv \mathfrak{Y}$ und 2): "Aus den vorangegebenen Sätzen der Form $X_i \equiv Y_i$ kann der Satz $\vdash \mathfrak{X} = \mathfrak{Y}$ allein mit Benutzung von Satz 5 und die Eigenschaften der Gleichheit bediesen werden," sind äquivalent.*

Beweis: Der allgemeinste Satz der Form $\mathfrak{X} \equiv \mathfrak{Y}$, den wir nach der Bedeutung von $\equiv$ aus den vorangegebenen schliessen können, ist ein Spezialfall von Satz 7 mit Gleichheit durch Identität ersetzt. Aber Satz 7 gibt auch den allgemeinsten Satz der Form $\vdash \mathfrak{X} = \mathfrak{Y}$ welcher aus $\vdash X_i = Y_i$ durch Benutzung der Eigenschaften der Gleichheit bewiesen werden kann. Damit wird die Behauptung bewiesen.

Festsetzung 1. Unter

$$\vdash X_1 = X_2 = X_3 \cdots = X_n$$

verstehen wir

$$\vdash X_1 = X_2 \text{ und } \vdash X_2 = X_3 \text{ und } \cdots \text{ und } \vdash X_{n-1} = X_n,$$

(woraus insbesondere folgt $\vdash X_1 = X_n$. Die Eigenschaften werden der Gleichheit hiernach im allgemeinen nicht besonders erwähnt).

SATZ 9. *Wenn X ein Etwas ist*

$$\vdash IX = X.$$

Beweis: $\vdash IX = WKX$ (C4. Def. 3, Satz 4),
$= KXX$ (Regel W),
$= X$ (Regel K).

Bemerkung: Anstatt B, C, W und K hat Schönfinkel K und S als Grundbegriffe gewählt. Dazu gehören die Regeln

$\vdash KXY = X$ für beliebige Etwase X, Y;
$\vdash SXYZ = XZ(YZ)$ für beliebige Etwase X, Y, Z,

woraus sich B, C, W und I wie folgt definieren lassen:

$$B \equiv S(KS)K, \quad C \equiv S(BBS)(KK), \quad W \equiv SS(SK), \quad I \equiv SKK.$$

In dieser Weise kann man wohl die Anzahl der Grundbegriffe und Regeln vermindern. Der Beweis der Eigenschaften der Gleichheit ist aber etwas schwieriger, und man muss vermutlich den obigen Satz 4 als Grundregel voraussetzen.

KAPITEL II. DIE LEHRE DER KOMBINATOREN.

A. EINLEITUNG.

Festsetzung 1. Unter einer *Kombination* von Etwasen $X_1, X_2 \cdots X_n$ verstehen wir ein Etwas, das sich aus $X_1 \cdots X_n$ durch Anwendung aufbauen lässt. Genauer definieren wir so:

1) Jedes X_i $(i = 1, 2, \cdots, n)$ ist eine Kombination von $X_1 \cdots X_n$.
2) Wenn X und Y Kombinationen von $X_1, \cdots X_n$ sind, so ist auch (XY) eine solche.

Festsetzung 2. Unter einem *Kombinator* verstehen wir eine Kombination von B, C, W und K.

In diesem Kapital untersuchen wir die Kombinatoren als solche. Das Hauptergebnis betrifft aber eine Verwandtschaft mit den Substitutionsprozessen. Es wird gezeigt, dass alle die in der gewöhnlichen Logik durch freie Variablen angedeuteten Verknüpfungen sich durch Kombinatoren definieren lassen, und zwar so, dass alle die Eigenschaften, die sie dort haben, auch aus unserem Grundgerüst ableitbar sind.

Um dieses Problem schärfer zu formulieren, betrachten wir diese Operationen und ihre Ausdrucksweise näher. Es soll genau erklärt werden, was sie sind und welche Eigenschaften in Betracht kommen.

Zunächst dienen natürlich die Variablen nur dazu, die Leerstelle in den verschiedenen Funktionen zu unterscheiden. Infolgedessen darf ich für sie eine besondere Bezeichnungsweise fordern, und zwar die folgende: Man beginnt mit gewissen Grundkonstanten und Grundfunktionen. In diesen sollen die Leerstelle mit den Zeichen $x_1, x_2 \cdots x_n$ konsekutiv (d. h. ohne Auslassungen)* numeriert werden, und zwar in einer Anordnung, die am Anfang beliebig zu wählen ist, aber danach festbleiben soll. Dann werden die folgenden Operationen für die Erschaffung neuer Funktionen und Aussagen erlaubt.

---

* [Oder genauer, wenn x_e als Argument vorkommt, so soll auch x_k für $k < e$ vorkommen. (P. Bernays)].

† In der gewöhnlichen Theorie sagt man zuweilen, dass ese zwischen $\phi(x, y)$ und $\phi(y, x)$ keinen Unterschied gebe, weil es nicht darauf ankomme, wie man die Variablen bezeichne. Aber wenn man eine Funktion $\phi(x, y)$ hat, so kann man andere Funktionen $\psi(x, y)$, $\chi(x)$ definieren, sodass

$$\psi(x, y) \equiv_{xy} \phi(y, x) \qquad \chi(x) \equiv_{x} \phi(x, x).$$

Dann ist der Inhalt meiner Forderung dieser: wenn wir z. B. ϕ mit $\phi(x_1, x_2)$ bezeichnen, dann werden ψ und χ mit $\phi(x_2, x_1)$ bzw. $\phi(x_1, x_1)$ bezeichnet. Natürlich haben wir hier im allgemeinen drei verschiedene Funktionen.

1) *Umwandlung* einer Funktion in eine zweite, die sich von der ersten nur darin unterscheidet, dass die Variablen anders numeriert sind. Diese neue Numerierung soll die obigen Bedingungen der Konsekutivität erfüllen, doch ist es erlaubt, Leerstellen, die ursprünglich verschiedene Zeichen trugen, dasselbe Zeichen zu geben, aber nicht umgekehrt. Z. B. aus $\phi(x_1, x_2)$ werden $\phi(x_2, x_1)$ und $\phi(x_1, x_1)$ durch Umwandlungen geschaffen usw. Die so umgeformten Funktionen, die verschiedene Bezeichnungen haben, sind als ganz voneinander und von der Ursprünglichen verschiedene Funktionen anzusehen.†

2) Einsetzung einer Konstanten a für eine Variable x_k in einer Funktion von n Variablen, wo $n \geqq k$. Dadurch wird eine Funktion von $n-1$ Variablen (bzw., wenn $n=1$ ist, eine Konstante) geformt. Dies soll in der folgenden Weise bezeichnet werden: für x_k in der ursprünglichen Funktion soll immer a auftreten, für x_i mit $i > k$ soll x_{i-1} auftreten, während die x_i mit $i < k$ unverändert bleiben sollen.

3) Zusammensetzung einer Funktion von n Variablen mit einer von m Variablen, durch Einsetzung von dieser *als Funktion von m Variablen* für x_k ($k \leqq n$) in jene. Die so gestaltete Funktion soll in der folgenden Weise bezeichnet werden: Zunächst setzt man in die zweite Funktion y_i für x_i ein, dann setzt man den so umgeformten Ausdruck als ein Ganzes in die erste Funktion in alle die Stellen ein, wo x_k da erscheint, und endlich formt man den resultierenden Ausdruck so um, dass x_i für $i < k$ unverändert bleibt, y_i in x_{k+i-1} übergeht, und x_i für $i > k$ in x_{i+m-1} übergeht. Z. B. durch Einsetzung von $\psi(x_2, x_1, x_1, x_3)$ für x_2 in $\phi(x_3, x_2, x_1, x_2, x_4, x_5, x_6)$ hat man $\phi(x_5, \psi(x_3, x_2, x_2, x_4), x_1, \psi(x_3, x_2, x_2, x_4), x_6, x_7, x_8)$. (Wenn man eine Konstante als eine Funktion von 0 Variablen ansieht, so ist der Fall (2) als Spezialfall im Falle (3) eingeschlossen).

Diese Operationen sind die, mit denen wir uns in diesem Kapitel zu beschäftigen haben. Die Gesamtheit der Ausdrücke, die sie erzeugen, hat die Eigenschaft, dass jede Funktion und jede Aussage, die aus den Grundfunktionen und Grundkonstanten gebildet werden kann, mit einem und nur einem Ausdruck bezeichnet wird. Es kann aber geschehen, dass derselbe Ausdruck in mehreren Weisen durch diese Operationen erzielt wird, also dass ganz verschiedene Operationsprozesse in dem Sinne äquivalent sind, dass sie dasselbe Etwas liefern. Diese Äquivalenzen sind die Eigenschaften, die hier in Betracht kommen.

Diese Ausdrücke lassen sich nun in unsere Schreibweise umformen, wenn wir nur die Variablen behandeln, als ob sie Etwase wären. In der Tat geht ein Ausdruck der Form $f(u_1, u_2, \cdots u_n)$ wegen der Ausdeutung der Anwen-

dung (als Schönfinkelscher Funktionsbegriff) in $(fu_1u_2\cdots u_n)$ über. Z. B. wird der oben (in 3) betrachtete Ausdruck nach der Umformung

$$(\phi x_5(\psi x_3x_2x_2x_4)x_1(\psi x_3x_2x_2x_4)x_6x_7x_8).$$

Die neuen Ausdrücke sind also Kombinationen von gewissen Grundfunktionen, Grundkonstanten und Variablen. Diese Kombinationen lassen sich aber ferner aus Kombinationen von lauter Variablen erschaffen und zwar dadurch, dass man in eine der letzten für x_1 eine Grundfunktion nach dem obigen Prozess 2) einsetzt, dann für x_2 eine Grundkonstante oder Grundfunktion einsetzt, u. s. w., bis alle die im betreffenden Ausdruck erscheinenden Grundgegenstände eingesetzt sind. Z. B. wird das oben Geschriebene aus

$$(x_1x_7(x_2x_5x_4x_4x_6)x_3(x_2x_5x_4x_4x_6)x_8x_9x_{10})$$

durch Einsetzung von ϕ statt x_1 und ψ statt x_2 erzielt.

Diese letzte Bemerkung läuft darauf hinaus, dass die Kombinationen lauter Variablen Operatoren sind, die aus den gegebenen Funktionen und Konstanten alle möglichen abgeleiteten Funktionen und Konstanten durch Einsetzung, und zwar die Schönfinkelsche,* hervorrufen. In diesem Kapitel zeige ich, dass diese Operatoren nichts wesentlich anderes als eine bestimmte Klasse von Kombinatoren sind. In der Tat setze ich zuerst fest, dass ein Kombinator Y eine Kombination der Variablen $x_1, x_2\cdots x_n$, nämlich X dann und nur dann *darstellt*, wenn es formal,—d. h. durch Behandlung der Variablen als Etwase ohne besondere Eigenschaften—folgt, das

$$\vdash Yx_1x_2\cdots x_n = X,$$

Dann beweise ich die folgenden Hauptsätze:

I. *Wenn ein Kombinator Y eine Kombination von $x_1x_2\cdots x_n$ darstellt, so stellt er nur eine dar* (bewiesen in C 1).

II. *Zu jeder Kombination lauter Variabler gibt es mindestens einen Kombinator Y, der sie darstellt* (bewiesen in E 1).

III. *Wenn zwei Kombinatoren Y_1 und Y_2 dieselbe Kombination lauter Variabler darstellen, dann folgt es ohne Gebrauch von Variablen, dass* $\vdash Y_1 = Y_2$ (bewiesen in E 3).†

Aus diesen drei Hauptsätzen folgt leicht, was ich beweisen will. Denn zunächst wird jeder Ausdruck eindeutig in der Form $Yu_1u_2\cdots u_n$, wo $u_1, u_2\cdots u_n$ die in dem Ausdruck erscheinenden Grundgegenstände sind,

---

* Vgl. oben unter I A.

† Mit der unwichtigen Beschränkung dass Y_1 und Y_2 eigentlich (s. unten II E) sind.

dargestellt, und weiter lässt sich jeder Substitutionsprozess so durch Kombinatoren definieren, dass aus den Ausdrücken der Form $Yu_1u_2\cdots u_n$ immer neue Ausdrücke derselben Form erzeugt werden.

Dieses Resultat ist freilich nur ein Spezialergebnis der drei Hauptsätze, welche eine gewisse Art von Isomorphismus zwischen Kombinatoren und Kombinationen überhaupt aussagen. Mit dem Kombinator K kommen Kombinationen mit Auslassungen in Betracht. Wir könnten diese ausschliessen, und nur Kombinatoren, die Kombinationen von B, C, W und I sind, betrachten; aber die allgemeineren Sätze können fast so leicht bewiesen werden als die besonderen, und also habe ich den K behalten. Die Betrachtungen, die den K speziell betreffen, sind unten zuweilen nicht so ausführlich angegeben, wie die anderen.

Eine letzte Bemerkung kommt dazu. In der eben dargestellten Theorie wurden bisher Funktionen von verschiedenen Anzahlen von Variablen unterschieden; dagegen ist jedes Etwas wegen Reg. E eine Funktion beliebig vieler Variablen. Infolgedessen hat man mit jedem Etwas nicht bloss eine bestimmte endliche Anzahl von Leerstellen zu assoziieren, sondern eine unendliche Folge von Leerstellen. Durch irgendeinen der obigen Prozesse wird aber nur eine endliche Anzahl von Leerstellen gestört, also haben diese Folgen einen besonderen Charakter. Dies erklärt die Tatsache, dass von Abschnitt C ab Folgen einer bestimmten Art eine grosse Rolle spielen.

B. Die Grundlegenden Definitionen und Sätze.

§ 1. *Die B Sequenz.*

Def. 1. $B_1 \equiv B$; $\quad B_{n+1} \equiv BBB_n$, $\quad (n = 1, 2, 3, \cdots)$.
Def. 2. $B_0 \equiv I$.

Satz 1. *Wenn X, Y und Z Etwase sind, dann*

$$\vdash B_{n+1}XYZ = B_nX(YZ), \qquad (n = 0, 1, 2, \cdots).$$

Beweis: Für $n = 0$ folgt der Satz aus Def. 1 und 2, Reg. B, und I D Satz 9.
Es sei $n > 0$;

$$\begin{aligned} \vdash B_{n+1}XYZ &= BBB_nXYZ && \text{(Def. 1, I D Satz 6)},\\ &= B(B_nX)YZ && (\text{Reg.}\, B),\\ &= B_nX(YZ), \quad \text{w. z. b. w.} && (\text{Reg.}\, B).\end{aligned}$$

SATZ 2. *Wenn X, Y, Z_1, Z_2, $\cdots$, Z_m Etwase sind, dann*

$$\vdash B_{n+m}XYZ_1Z_2\cdots Z_m = B_nX(YZ_1Z_2\cdots Z_m), \qquad (m=0,1,2,\cdots;\ n=0,1,2,\cdots).$$

Beweis: Für $m=0$ klar. Für $m=1$ folgt der Satz aus Satz 1.

Angenommen, der Satz sei für $m=k$ bewiesen. Dann wird er für $m=k+1$ bewiesen, wie folgt: aus Satz 1 folgt

$$\vdash B_{n+k+1}XYZ_1\cdots Z_kZ_{k+1} = B_{n+k}X(YZ_1)Z_2Z_3\cdots Z_kZ_{k+1}, \qquad (n=0,1,2\cdots).$$

Nun in den vorliegenden Satz für $m=k$ setzen wir (YZ) für Y, und Z_{i+1} für Z_i $(i=1,2,\cdots,k)$. Dann

$$\vdash B_{n+k}X(YZ_1)Z_2Z_3\cdots Z_kZ_{k+1} = B_nX(YZ_1Z_2\cdots Z_kZ_{k+1}), \qquad (n=0,1,2,\cdots).$$

Aus den letzten zwei Formeln folgt der Satz für $m=k+1$. Durch Wiederholung dieses Prozesses wird der Satz für ein beliebiges m bewiesen.

SATZ 3. *Sind X, Y, Z_1, Z_2, $\cdots$, Z_m Etwase, dann*

$$\vdash B_mXYZ_1Z_2\cdots Z_m = X(YZ_1Z_2\cdots Z_m), \quad (m=0,1,2,3,\cdots).$$

Beweis: Man setzt $n=0$ in Satz 2.

SATZ 4. *Wenn eine Reihe von Etwasen X_1, X_2, X_3, $\cdots$, einer Rekursionsformel, näm.,*

$$\vdash X_{n+1} = BX_n \qquad (n=1,2,3,\cdots),$$

erfüllen, dann

$$\vdash X_{n+m} = B_mX_n, \qquad (m=0,1,2,\cdots;\ n=1,2,3,\cdots).$$

Beweis: Für $m=0$ oder $m=1$ klar.
Nun sei der Satz für $m=k$ augenommen. Dann folgt

$$\begin{aligned}
\vdash X_{n+k+1} &= BX_{n+k} && \text{(Hp.)},\\
&= B(B_kX_n) && \text{(nach diesem Satz für } m=k),\\
&= BBB_kX_n && \text{(Reg. } B),\\
&= B_{k+1}X_n && \text{(Def. 1)}.
\end{aligned}$$

Also kann der Satz für ein beliebiges m bewiesen werden.

SATZ 5. *Ist X ein Etwas, dann*

$$\vdash B_m(B_nX) = B_{m+n}X, \qquad (m,n=0,1,2,3,\cdots).$$

Beweis: Für $n = 0$ klar, weil $\vdash B_0 X = X$. Sonst setzt man:

$$X_n \equiv B_n X, \qquad (n = 1, 2, 3, \cdots),$$

dann

$$\begin{aligned} \vdash X_{n+1} &= B_{n+1}X = BBB_nX && \text{(Def. 1)}, \\ &= B(B_nX) = BX_n && \text{(Reg. } B). \end{aligned}$$

Der Satz folgt dann aus Satz 4.

§ 2. *Die Identitätsfunktion.*

SATZ 1. $\vdash B_m I = I, \qquad (m = 0, 1, 2, \cdots).$

Beweis: In § 1 Satz 4 setzen wir $X_n \equiv I$. Dann

$$\vdash X_{n+1} = BX_n \qquad \text{(Ax. } I_2).$$

Daher folgt der Satz aus § 1 Satz 4.

SATZ 2. *Wenn X ein Etwas ist, dann*

$$\vdash B_m I X = X, \qquad (m = 1, 2, 3, \cdots).$$

Beweis: $\vdash B_m I X = IX$ (Satz 1),
$= X,$ (I D Satz 9).

SATZ 3. $\vdash CBI = I.$

Beweis:
$$\begin{aligned} \vdash CBI &= B(BI)I && \text{(Ax. } I_1), \\ &= BII = II && \text{(Ax. } I_2), \\ &= I. && \text{(I D Satz 9)}. \end{aligned}$$

SATZ 4. *Wenn X ein Etwas ist, dann gilt* $\vdash BXI = X.$

Beweis:
$$\begin{aligned} \vdash BXI &= CBIX && \text{(Reg. } C), \\ &= IX && \text{(Satz 3)}, \\ &= X && \text{(I D Satz 9)}. \end{aligned}$$

SATZ 5. *Def.* 1 *von* § 1 *gilt auch, wenn* $n = 0$, *d. h.*

$$\vdash B_1 = BBB_0.$$

Beweis: Klar aus Satz 4 und § 1, Def. 2.

§ 3. *Die C, W und K Sequenzen.*

Def. 1. $C_1 \equiv C;\quad C_{n+1} \equiv BC_n, \qquad (n = 1, 2, 3, \cdots).$

SATZ 1. $\vdash B_m C_n = C_{m+n}, \qquad (m = 0, 1, 2, \cdots;\ n = 1, 2, 3, \cdots).$

Beweis: Folgt unmittelbar aus Def. 1 und § 1 Satz 4.

Satz 2. *Wenn $X_0, X_1, X_2, \cdots, X_n, X_{n+1}$, Etwase sind, dann*

$$\vdash C_n X_0 X_1 X_2 \cdots X_n X_{n+1} = X_0 X_1 X_2 \cdots X_{n-1} X_{n+1} X_n.$$

Beweis:
$$\begin{aligned} \vdash C_n X_0 X_1 \cdots X_{n-1} X_n X_{n+1} &= B_{n-1} C_1 X_0 X_1 \cdots X_{n+1} && \text{(Satz 1)},\\ &= C_1 (X_0 X_1 X_2 \cdots X_{n-1}) X_n X_{n+1} && \text{(§ 1 Satz 3)},\\ &= X_0 X_1 X_2 \cdots X_{n-1} X_{n+1} X_n, && \text{(Reg. } C). \end{aligned}$$

Def. 2. $W_1 \equiv W$; $\quad W_{n+1} \equiv B W_n$, $\quad (n = 1, 2, 3, \cdots)$.

Satz 3. $\vdash B_m W_n = W_{m+n}$, $\quad (m = 0, 1, 2, \cdots; n = 1, 2, 3, \cdots)$.

Beweis: Folgt unmittelbar aus Def. 2 und § 1, Satz 4.

Satz 4. *Wenn $X_0, X_1, X_2, \cdots X_n$ Etwase sind, so gilt*

$$\vdash W_n X_0 X_1 \cdots X_n = X_0 X_1 \cdots X_{n-1} X_n X_n.$$

Beweis: Wie der von Satz 2, mit Gebrauch von Reg. W anstatt Reg. C.

Def. 3. $K_1 \equiv K$; $\quad K_{n+1} \equiv B K_n$, $\quad (n = 1, 2, 3, \cdots)$.

Satz 5. $\vdash B_m K_n = K_{m+n}$, $\quad (m = 1, 2, 3, \cdots; n = 1, 2, 3, \cdots)$.

Satz 6. *Wenn $X_0, X_1, X_2, \cdots X_n$ Etwase sind, so gilt*

$$\vdash K_n X_0 X_1 X_2 \cdots X_n = X_0 X_1 X_2 \cdots X_{n-1}.$$

Beweis: Wie der von Satz 2 mit Gebrauch von Reg. K anstatt Reg. C.

§ 4. *Das zusammengesetzte Produkt.*

Def. 1. Wenn X und Y Etwase sind, $(X \cdot Y) \equiv BXY$.

Satz 1. *Sind X, Y, Z Etwase, dann* $\vdash (X \cdot Y) Z = X(YZ)$.

Beweis: Folgt aus Def. 1 und Reg. B.

Def. 2. Wenn X und Y Bezeichnungen für Etwase sind, und keine von den beiden von der Form $(U = V)$ oder $(U \cdot V)$ ist, so dürfen wir in einer Bezeichnung wie $X \cdot Y$ die äusseren Klammern von X und Y fortlassen.

Satz 2. *Sind X und Y Etwase, dann gilt*

$$\vdash B(X \cdot Y) = BX \cdot BY.$$

Beweis:

$$(1)\quad \vdash B(X\cdot Y) = B(BXY) \qquad \text{(Def. 1)},$$
$$= B_2BBXY \qquad \text{(§ 1, Satz 3)}.$$
$$(2)\quad \vdash BX\cdot BY = B(BX)(BY) \qquad \text{(Def. 1 und 2)},$$
$$= B_2X(BY) \qquad \text{(§ 1, Satz 5, } m=n=1\text{)},$$
$$= B_3XBY \qquad \text{(§ 1, Satz 2, } m=1,\ n=2\text{)},$$
$$= CB_3BXY \qquad \text{(Reg. } C\text{)}.$$

Nun aus § 1, Def. 1,

$$\vdash CB_3B = C(BB(BBB))B,$$

daher aus Ax. B

$$(3)\quad \vdash CB_3BXY = B(BB)BXY$$
$$= B_2BBXY \qquad \text{(§ 1, Satz 5)}.$$

Aus (1), (2), (3), wird der Satz bewiesen.

Satz 3. *Das Produkt* $(X\cdot Y)$ *ist assoziativ, d. h., wenn* X, Y, Z *Etwase sind, so gilt*

$$\vdash X\cdot(Y\cdot Z) = (X\cdot Y)\cdot Z.$$

Beweis:

$$\vdash X\cdot(Y\cdot Z) = B(X\cdot Y)Z \qquad \text{(Def. 1)},$$
$$= (BX\cdot BY)Z \qquad \text{(Satz 2)},$$
$$= BX(BYZ) \qquad \text{(Satz 1)},$$
$$= X\cdot(Y\cdot Z) \qquad \text{(Def. 1)}.$$

Satz 4. *Wenn* X *ein Etwas ist, so gelten*

1) $\vdash I\cdot X = X,$
2) $\vdash X\cdot I = X.$

Beweis: Folgt aus Def. 1 und § 2, Sätze 2 und 4.

Satz 5. $\vdash B_m\cdot B_n = B_{m+n}, \quad (m, n = 1, 2, 3, \cdots).$

Beweis: Für $m=1$ folgt aus Def. 1 und § 1, Def. 1. Ist nun der Satz für $m=k$ bewiesen, dann folgt

$$\vdash B_{k+1}\cdot B_n = (B\cdot B_k)\cdot B_n \qquad \text{(nach diesem Satze für } m=1\text{)},$$
$$= B\cdot(B_k\cdot B_n) \qquad \text{(Satz 3)},$$
$$= B\cdot B_{n+k} \qquad \text{(nach diesem Satze für } m=k\text{)},$$
$$= B_{n+k+1}, \qquad \text{(nach diesem Satze für } m=1\text{)}.$$

Dadurch wird der Satz für ein beliebiges m bewiesen.

SATZ 6. *Sind X und Y Etwase, dann gilt,*

$$\vdash B_m(X \cdot Y) = B_m X \cdot B_m Y, \qquad (m = 1, 2, 3, \cdots).$$

Beweis: Für $m = 1$ ist der Satz mit Satz 2 identisch. Nun sei der Satz für $m = k$ angenommen, dann wird er für $m = k + 1$ bewiesen, wie folgt:

$$\begin{aligned} \vdash B_{k+1}(X \cdot Y) &= B(B_k(X \cdot Y)) && (\S\, 1, \text{ Satz } 5), \\ &= B(B_k X \cdot B_k Y) && (\text{nach diesem Satze für } m = k) \\ &= B(B_k X) \cdot B(B_k Y) && (\text{Satz } 2), \\ &= B_{k+1} X \cdot B_{k+1} Y, && (\S\, 1, \text{ Satz } 5). \end{aligned}$$

SATZ 7. $\vdash B_m B_n \cdot B_m B_p = B_m B_{n+p}, \quad (m, n, n, = 0, 1, 2, \cdots).$

Beweis: Folgt aus Sätzen 4,* 5, 6.

Def. 3. Sind $X_1, X_2, \cdots, X_n$ Etwase, dann

$$(X_1 \cdot X_2 \cdot X_3 \cdots X_n) \equiv (\cdots(((X_1 \cdot X_2) \cdot X_3) \cdot X_4) \cdots X_n).$$

Festsetzung: Da ich die Assoziativität des Produkts $(X \cdot Y)$ bewiesen habe, brauche ich diese Tatsache nicht immer explizit hervortreten zu lassen. Im Gegenteil werde ich die Form der Def. 3 benutzen, und dann habe ich ein Recht, Klammern nach Belieben einzusetzen. Die Def. 3 setzt aber an sich die Assoziativität nicht voraus, und wird benutzt, wo die Assoziativität nicht behauptet ist (z. B. s. die nächste Nummer).

§ 5. *Die Axiome in der neuen Darstellungsweise.*

SATZ. *Wenn man die Definitionen von §§ 1-4 berücksichtigt, nehmen die kombinatorischen Axiome die folgende Gestalt an:*

a. *Kommutative Axiome.*

Ax. B. $\vdash CB_3B = BB \cdot B.$
Ax. C. $\vdash CB_3C = BC \cdot B_2.$
Ax. W. $\vdash CB_2W = BW \cdot B_2.$
Ax. K. $\vdash CB_2K = BK \cdot I.$
Ax. I_1. $\vdash CBI = BI \cdot I.$

b. *Transmutative Axiome.*

Ax. (BC). $\vdash B_1 \cdot C_1 = C_2 \cdot C_1 \cdot BB.$
Ax. (BW). $\vdash B_1 \cdot W_1 = W_2 \cdot W_1 \cdot C_2 \cdot B_2B \cdot B.$
Ax. (BK). $\vdash B_1 \cdot K_1 = K_1 \cdot K_1.$

---

* Satz 4 ist nur wenn $n = 0$ oder $p = 0$ notwendig.

Ax. $(CC)_1$. $\vdash C_1 \cdot C_1 = B_2 I$.
Ax. $(CC)_2$. $\vdash C_2 \cdot C_1 \cdot C_2 = C_1 \cdot C_2 \cdot C_1$.
Ax. (CW). $\vdash C_1 \cdot W_1 = W_2 \cdot C_1 \cdot C_2$.
Ax. (CK). $\vdash C_1 \cdot K_1 = K_2$.
Ax. (WC). $\vdash W_1 \cdot C_1 = W_1$.
Ax. (WW). $\vdash W_1 \cdot W_1 = W_1 \cdot W_2$.
Ax. (WK). $\vdash W_1 \cdot K_1 = BI$.

c. *Andere Axiome.*

Ax. I_2. $\vdash BI = I$.

Bemerkungen über die Bedeutung dieser Axiomen. Diese Axiomen haben einen Vertauschbarkeitscharakter, d. h., sie sagen aus, dass gewisse Ausdrücke in zwei Weisen gebildet werden können. Z. B. kann der Ausdruck

$$x_0 x_3 (x_1 x_2)$$

aus dem Ausdruck

$$x_0 x_1 x_2 x_3$$

in den folgenden zwei Weisen gebildet werden: 1), man setzt zunächst die Klammern ein, und führt dann eine Vertauschung aus, oder 2), man macht zuerst eine Vertauschung und dann eine Einsetzung von Klammern. Diese zwei Konstructionsweisen entsprechen den zwei Darstellungen

$$(B \cdot C) x_0 x_1 x_2 x_3$$
$$(C_2 \cdot C_1 \cdot BB) x_0 x_1 x_2 x_3$$

die nach Ax. (BC) gleich sind. Ebenso sagen die sämtlichen transmutativen Axiomen eine Vertauschbarkeit aus. Dass die kommutativen Axiomen in dieser Hinsicht nicht wesentlich verschieden sind, folgt aus II D 2, Satz 1 (s. unten).

Der Unterschied zwischen den kommutativen und transmutativen Axiomen liegt darin, dass die Kombinatoren, deren Vertauschbarkeit durch jene Axiomen ausgesagt wird, nicht miteinander übergreifen. Um dies zu erklären, setzen wir voraus, dass für einen Kombinator X gilt

$$\vdash X x_0 x_1 x_2 \cdots x_m = x_0 y_1 y_2 \cdots y_n \; *$$

wo $y_1, y_2, \cdots, y_n$ Kombinationen von $x_1, x_2, \cdots, x_m$ sind. Ist Y irgendein

* Diese Gleichung ist in dem Sinne zu verstehen, dass die Variablen wie beliebige Etwase betrachtet werden.

Kombinator der Form $B_m Z$, so ist es ersichtlich, dass Y auf Variablen wirkt, die durch X nicht gestört werden, und umgekehrt; also muss eine Gleichung der Form

$$B_m Z \cdot X = X \cdot B_n Z$$

inhaltlich richtig sein. Diese Gleichung sagt aber eine Vertauschbarkeit derjenigen Art aus, welche aus den kommutativen Axiomen folgt. Die transmutativen Axiomen haben aber den entgegengesetzten Charakter. Z. B. wirken die zwei Kombinatoren B und C, die auf der linken Seite des *Ax.* (BC) erscheinen, auf dieselben Variablen, näm. x_1 und x_2.

Das Ax I_2 ist ein Axiom ganz anderer Art. Seine wesentliche Bedeutung ist dass es uns ermöglicht, die Gleichheit von Kombinatoren zu beweisen, die im Sinne von II C 1 (unten) derselben Folge von lauter Variablen, aber mit verschiedenen Ordnungen entsprechen—was sonst unmöglich wäre (II C 1, Satz 6). D. h., dieses Axiom sagt aus, dass ein Kombinator unabhängig davon zu verstehen ist, wie viele Variablen hinzuzufügen sind, damit eine Reduktion auf eine Kombination von lauter Variablen sich vollzieht.

A Property of Unbounded Continua, with Applications.*

By W. A. Wilson.

1. A great part of the theory of continua is valid only if the continua under discussion are bounded. This is primarily due, of course, to the fact that for unbounded sets it is not necessarily true that every infinite set has at least one limiting point. Hence such fundamental theorems as the Cantor divisor theorem and the existence of various kinds of irreducible continua fail.† However, many theorems which are true only for bounded continua are readily generalized for the unbounded case by the imposition of suitable restrictions. The favorite tool used for such extensions has been the method of inversion.‡

It is the purpose of this article to call attention to a property of unbounded continua (§ 2) which is useful in investigations of this kind and to give applications of it to irreducible continua and the separation of the plane. It is possible that this property has been noted before by other students of the subject (Cf. the article by Knaster and Kuratowski referred to above and the article by Szymanski referred to in § 5), but, to judge by the literature, it has been discarded as of no importance. The material of §§ 2-5 is valid in any metric space for which the Bolzano-Weierstrass property that every bounded infinite set has at least one limiting point holds; for the sake of brevity such a space will be called a W-space.

2. Theorem. *Let $\{X_i\}$ be a sequence of continua lying in a W-space and let the upper closed limit K of the sequence be a non-void. Let each X_i be unbounded or let the diameter of X_i increase indefinitely with i. Then K is an unbounded continuum or a closed set each component of which is an unbounded continuum.*

* Presented to the American Mathematical Society, December, 1929.

† A conspicuous exception is that in a metric connected space, which is everywhere locally connected and in which every bounded infinite set has one or more limiting points, every closed cut between two points contains an irreducible cut between the points.

‡ The principal properties of inversion may be found systematically worked out in the following articles: C. Kuratowski, "Sur la méthode d'inversion dans l'Analysis Situs," *Fundamenta Mathematicae*, Vol. 4, pp. 151-163; and B. Knaster and C. Kuratowski, "Sur les continus non-bornés," *Ibid.*, Vol. 5, pp. 23-58.

Proof. Let us assume the opposite, namely that K has a bounded component A. Let $\epsilon > 0$, S denote the set of points whose distances from A are less than ϵ, and F be the frontier of S. By the definition of upper closed limit there is for each point a of A a partial sequence $\{i'\}$ of $\{i\}$ and a set of points $\{a_{i'}\}$ such that each $a_{i'}$ lies in $X_{i'}$ and $a_{i'} \to a$. For i' large enough every $a_{i'}$ lies in S and $X_{i'}$ contains points not in $\bar{S}$. Since $\bar{S}$ is bounded, $X_{i'}$ contains a sub-continuum $Y_{i'}$ contained in $\bar{S}$ and joining $a_{i'}$ to a point $y_{i'}$ of F. Since the union of the continua $\{Y_{i'}\}$ is bounded, the upper closed limit of the sequence $\{Y_{i'}\}$ is non-void and is a continuum joining a to a point y on F. But, as $Y \subset K$, it follows that $Y \subset A$. This is a contradiction, for $A \subset S$ and y lies in F.

COROLLARY. *Let $\{X_i\}$ be a monotone decreasing sequence of unbounded continua lying in a W-space and let D be their divisor. If D is not void, it is an unbounded continuum or a closed set of which each component is an unbounded continuum.*

Note 1. If in the above theorem we impose no restrictions on the diameters of the continua $\{X_i\}$, but require that the lower closed limit of the sequence be non-void, we find that K is either a continuum or a closed set of which each component is an unbounded continuum. The proof is a modification of that given above.

Note 2. A consequence of the above corollary is that in C. Kuratowski's decomposition of the irreducible continuum into *tranches*,* each tranche is a continuum or a closed set of which each component is an unbounded continuum.

3. A generalization. It is convenient to say that an unbounded continuum containing points of a set α joins α to ∞. If a continuum K joins a set α to ∞ and no proper sub-continuum has this property, K is called *irreducible between α and ∞*. If a continuum is irreducible between two sets α and β, and also between α and ∞, we say that it is irreducible between α and $\beta + \infty$. The following examples, all of which are sets in the Cartesian plane, make this notion clearer.

Example I. Let H be the set defined by $0 < x \leqq 1$ and $y = 1/x$. Then H is a continuum irreducible between $(1, 1)$ and ∞.

Example II. Let K be the set defined by $0 < x \leqq 1$ and $y = 1/x \sin^2(\pi/x)$.

* "Théorie des continus irréductibles entre deux points II," *Fundamenta Mathematicae*, Vol. 10, p. 254.

Then K is not a continuum, although it is homeomorphic with the set H in Ex. I.

Example III. Let $P=\bar{K}$, where K is the set in Ex. II. Let $\alpha=(1,0)$ and β be any set for which $x=0$ and $y\geqq 0$. Then P is a continuum irreducible between α and $\beta+\infty$.

Example IV. Let P be as in Ex. III and Q be the set symmetrical to P with respect to the origin. Set $M=P+Q$, $\alpha=(1,0)$ and $\beta=(-1,0)$. Then M is a continuum irreducible between α and β, but not between α or β and ∞.

Example V. Let M be defined thus: for $x=0$ and $x=1$, $y\geqq 0$; for $x=1/n$ and $x=1-1/n$, $0\leqq y\leqq n$; for $1/n<x<1-1/n$, $y=n$; for $1/(2n+1)<x<1/2n$ and $1-1/(2n+1)<x<1-1/(2n+2)$, $y=0$. (Here $n=1, 2, \cdots$) Then the continuum M contains no continuum irreducible between a set α on the line $x=0$ and a set β on the line $x=1$; but, if γ is any bounded closed sub-set which has no points on the lines $x=0$ and $x=1$, M contains a sub-continuum irreducible between γ and $\alpha+\beta+\infty$.

It is a simple matter to show that, if K is a continuum irreducible between a point a and ∞, and is everywhere locally connected, then K is the homeomorphic image of a ray. Furthermore, if K is simultaneously irreducible between the sets α and β, between α and ∞, and between β and ∞, it is indecomposable.

4. THEOREM. *Let the unbounded continuum M lie in a W-space and contain points of the bounded closed set α. Then M contains a continuum irreducible between α and ∞.*

Proof. Brouwer's theorem for establishing the existence of closed sets irreducible with respect to a property is valid here. For, if $\{K_i\}$ denotes any descending sequence of sub-sets of M, each of which contains points of α and has only unbounded components, and D is the divisor of this sequence, we have these results: $D\cdot\alpha\neq 0$, since each set $\alpha\cdot K_i$ is a bounded closed set; D need not be a continuum, but the components of D are all unbounded continua by § 2 and at least one of them contains points of α.

Note. This theorem need not be valid if α is unbounded or not closed. For example, let M be the positive half of the x-axis and $\alpha=\{n\}$, where $n=1, 2, \cdots$.

5. THEOREM. *Let M be an unbounded continuum in a W-space which contains points of the bounded closed sets α and β. Let M have a sub-*

continuum C which contains points of α and β, but does not contain two sub-continua without common points which join α and β, respectively, to ∞. Then there is a sub-continuum of M irreducible between α and β.

Proof. The method of § 4 applies. For, if $\{K_i\}$ is a descending sequence of sub-continua of C which contain points of α and β, and D is the divisor of this sequence, D contains a continuum joining α to β, as otherwise, by § 2, it would have at least two unbounded components, one containing points of α, and the other points of β.

Note. The condition given in the above theorem is obviously not necessary. It is closely related to a lemma proved by P. Szymanski † by means of inversion, which may be stated thus: Let A and B be closed sets and A be bounded; let S be a closed set such that $S \cdot A \neq 0 \neq S \cdot B$ and there is no decomposition of S into two closed sets M and N such that $M \cdot N = M \cdot B = N \cdot A = 0$. If S contains no sub-continuum joining A to B, then S contains an unbounded continuum K such that $K \cdot A \neq 0$.

6. Some properties of inversion. If A is a set in a Euclidean space and v is a center of inversion not on A, the inverse of A will be denoted by A^*; whence $A^{**} = A$. This notation and the properties of inversion obtained in the articles of Knaster and Kuratowski referred to in § 1 will be freely used in the following sections.

The following theorems regarding irreducible continua in Euclidean spaces are easily demonstrated:

I. *Let α be a bounded closed set, M a continuum irreducible between α and ∞, and v a point not on $M + \alpha$. Then $M^* + v$ is a continuum irreducible between α^* and v.*

II. *Let the bounded continuum M be irreducible between the closed sub-sets α and β, and v be a point of β. Then $(M - v)^*$ is an unbounded continuum irreducible between α^* and $(\beta - v)^* + \infty$.*

III. *Let the bounded continuum M be irreducible between the closed sub-sets α and β, and v be a point of $M - (\alpha + \beta)$. If $M - v$ is connected, $(M - v)^*$ is an unbounded continuum irreducible between α^* and β^*. If $M - v$ is not connected, $(M - v)^*$ is the sum of two continua P^* and Q^* without common points, which are irreducible between α^* and β^*, respectively, and ∞.*

These theorems lead at once to the following definitions. Let the set M

† "La somme de deux continus irréductibles," *Fundamenta Mathematicae*, Vol. 11, p. 7.

have points in common with each of the sets α and β. Then M is called *unstable with respect to α and β,* if it contains two unbounded continua P and Q such that $P \cdot Q = 0$, $P \cdot \alpha \neq 0$, and $Q \cdot B \neq 0$. Otherwise M is *stable* with respect to the two sets. This distinction is of especial importance in the inversion of irreducible continua, as is shown by the following theorem, which is a modification of one by P. Szymanski (*loc. cit.*, p. 7).

IV. *Let the unbounded continuum M be irreducible between the closed sub-sets α and β. Let v be a point not on M. For $M^* + v$ to be irreducible between α^* and β^* it is necessary and sufficient that M be stable with respect to α and β.*

7. Separation of the plane. Since the property of being an irreducible cut of the plane is an invariant under inversion, the theorems of the previous section enable one to extend to unbounded continua certain theorems regarding frontier sets in the plane.*

THEOREM. *Let $n \geqq 2$ be an integer and let the unbounded continuum F be the union of two continua H and K having one of the following sets of properties:* (1) *$H \cdot K$ is the sum of $n-1$ bounded closed sets $\{\alpha_i\}$ and H and K are irreducible between each pair of the sets $\{\alpha_i\}$ and between each α_i and ∞;* (2) *$H \cdot K$ is the sum of n bounded closed sets $\{\alpha_i\}$, H and K are irreducible between each pair of these, H is bounded, and K is stable with respect to each pair; or* (3) *$H \cdot K$ is the sum of $n-1$ bounded closed sets $\{\alpha_i\}$ and another closed set β, and both H and K are irreducible between each pair of sets $\{\alpha_i\}$ and between each α_i and $\beta + \infty$. Then F is the frontier of at least n components of its complement.*

Proof. Invert the plane with respect to a point v not on F. In Case 1, H and K are stable with respect to each pair of closed sets $\{\alpha_i\}$; hence by § 6, Theorems I and IV, $H^* + v$ and $K^* + v$ are irreducible between each pair of the n closed sets v and $\{\alpha_i^*\}$, while $(H^* + v)\cdot(K^* + v) = v + \sum_1^{n-1}\alpha_i^*$. In Case 2 for like reasons H^* and $K^* + v$ are irreducible between each pair of the n closed sets $\{\alpha_i^*\}$; and $H^* \cdot (K^* + v) = \sum_1^n \alpha_i^*$. In Case 3 both H and K are stable with respect to each pair of the n sets β and $\{\alpha_i\}$. Hence $H^* + v$ and $K^* + v$ are irreducible between each pair of the n sets $\{\alpha_i^*\}$ and $\beta^* + v$, while $(H^* + v)\cdot(K^* + v) = H^* \cdot K^* + v = \beta + v + \sum_1^{n-1}\alpha_i^*$.

Thus in every case $F^* + v$ is the union of two bounded continua whose divisor is the sum of n closed sets between each pair of which both continua are irreducible. Hence $F^* + v$ is the frontier of at least n components of its

* W. A. Wilson, "On Bounded Regular Frontiers," *Bulletin of the American Mathematical Society*, Vol. 34, p. 86, and C. Kuratowski, "Sur la séparation d'ensembles situés sur le plan," *Fundamenta Mathematicae*, Vol. 12, p. 235.

complement by the theorem cited in the foot-note above, and by inversion this is also true of F.

THEOREM. *Conversely, let F be a decomposable unbounded plane continuum which is the frontier of n ($\geqq 2$) components of its complement. Then F satisfies one of the sets of conditions of the previous theorem.*

Proof. Invert with respect to a point v not on F. Then $F^* + v$ is a decomposable bounded regular frontier which is the union of two continua H' and K' such that $H' \cdot K'$ is the sum of n closed sets $\{\alpha_i'\}$ between each pair of which both H' and K' are irreducible. There are four cases.

I. Some α_i', say α_n', is the point v. Invert with respect to v again, setting $H = (H' - v)^*$ and $K = (K' - v)^*$. Then $H \cdot K = \sum_1^{n-1} \alpha_i$, where $\alpha_i = \alpha_i'^*$ and each α_i is bounded. By § 6, Theorem II, both H and K are irreducible between each pair of the sets $\{\alpha_i\}$, $i \leqq n-1$, and between each α_i and ∞. Thus we have Case 1 of the previous theorem.

II. v is a point of some α_i', say of α_n', and $\alpha_n' - v \neq 0$. Proceed as in the first case, setting $\alpha_i = \alpha_i'^*$ for $i < n$, and $\alpha_n = (\alpha_n' - v)^*$. Since α_n' is closed, so is α_n. Here $H \cdot K = \sum_1^n \alpha_i$, and H and K are irreducible between each pair of the sets $\{\alpha_i\}$ for $i < n$, and between each α_i for $i < n$ and $\alpha_n + \infty$. Thus we have Case 3 of the previous theorem.

III. The point v lies in $H' - H' \cdot K'$ or in $K' - H' \cdot K'$, say the latter, and does not disconnect K'. Invert again, setting $H = H'^*$ and $K = (K' - v)^*$. Then each $\alpha_i = \alpha_i'^*$ is a bounded set and both H and K are irreducible between each pair of the sets $\{\alpha_i\}$. As $H' \cdot v = 0$, H is bounded. As $K' = K^* + v$, K is stable with respect to each pair of the sets $\{\alpha_i\}$ by § 6, Theorem IV. Thus we have Case 2 of the previous theorem.

IV. The point v lies in $K' - H' \cdot K'$, as above, and disconnects K'. Then K' is decomposable, and this can only happen when $n = 2$. This gives $K' = P' + Q'$, where $P' \cdot Q' = v$ and P' and Q' are continua irreducible between α_1' and α_2', respectively, and v. Then $F^* + v = P' + (H' + Q')$, $P' \cdot (H' + Q') = \alpha_1' + v$, and the continua P' and $H' + Q'$ are irreducible between α_1' and v. Invert again, setting $P = P'^*$, $H + Q = (H' + Q')^*$, and $\alpha_1 = \alpha_1'^*$. Then $P \cdot (H + Q) = \alpha_1$ and both P and $H + Q$ are continua irreducible between α_1 and ∞ by § 6, Theorem II. Thus we have Case 1 of the previous theorem once more.

Remarks. At first sight the three sets of conditions satisfied by the frontier in the above theorems appear to be unnecessarily complicated. But examples can be constructed showing the effectiveness of each set of conditions; it does not seem worth while, however, to give a full set of these, since for the case that $n = 2$ we can find examples for all three sets by merely modifying the familiar graph of $y = \sin 1/x$.

YALE UNIVERSITY, NEW HAVEN, CONN.

Centers of Symmetry in Analysis Situs.*

By Harry Merrill Gehman.

1. *Introduction.* The results of this paper were suggested by the fact that a simple continuous arc is "symmetrical" with respect to any one of its interior points in the sense that the complement of an interior point consists of two mutually separated sets which are homeomorphic † in such a way that the homeomorphism between them can be extended to the center of symmetry.

A new characterization of a simple continuous arc is given in Part 5, using the type of center of symmetry which I have called a "symmetrical cut point."

2. *Symmetry in a Euclidean n-Space.* Thruout the present section, let M denote a point set lying in an n-dimensional Euclidean space S. We shall say that P is a *center of symmetry of M* (in the sense of Analysis Situs), if there exists in S an n-dimensional rectangular coördinate system T having P as origin, which has the property that if $A=(x_1, x_2, \cdots, x_n)$ is a point of M, then $A_T=(-x_1, -x_2, \cdots, -x_n)$ is also a point of M. It will be convenient to refer to A and A_T as a *pair of corresponding points.* If P is itself a point of M, the points P and P_T are identical, but in all other cases, the points of a pair of corresponding points are distinct.

From this definition of center of symmetry, a number of obvious theorems can be deduced, of which the two following are of some interest.

Theorem 1. *If P is a center of symmetry of the point set M, and if P is neither a point nor a limit point of M, then every point within an n-dimensional sphere enclosing P but enclosing no point of M, is also a center of symmetry of M.*

This is true because a new coördinate system T' can be set up, which effects a change of origin from P to any other given point within the sphere, but does not affect the coördinates of points on and outside the sphere.

Theorem 2. *If P is a center of symmetry of M and if $(M+P)-P$*

* Presented to the American Mathematical Society at Ann Arbor, Michigan, on November 29, 1929.

† The sets M_1 and M_2 are said to be *homeomorphic* if there exists a continuous $(1-1)$ correspondence Π such that $\Pi(M_1)=M_2$.

is expressed in any way as the sum of two sets M_1 and M_2 which are such that neither set contains both points of any pair of corresponding points, then there exists a continuous $(1-1)$ correspondence Π having these properties: $\Pi(S)=S$, $\Pi(M)=M$, $\Pi(M_1)=M_2$, $\Pi(P)=P$, and $\Pi^{-1}(A)=\Pi(A)$, *for each point A of S.*

The proof of this theorem is obvious, since the correspondence may be defined as a reflection of space about P as center, with respect to the coördinate system T.

3. *Symmetry in a Point Set.* Thruout the remainder of this paper, we shall let M denote a point set which comprises the totality of points under consideration. That is, M is now thought of as a "space," and no properties are assumed of another space within which M may lie. Center of symmetry cannot now be defined thru the intermediation of a coördinate system, and we shall use instead certain properties of continuous $(1-1)$ correspondences as a basis for our definition of center of symmetry.

Definition. We shall say that the point P is a *center of symmetry of M,* if $M-P$ can be expressed as the sum of two mutually exclusive sets M_1 and M_2 which are such that there exists a continuous $(1-1)$ correspondence Π having the properties: $\Pi(M)=M$, $\Pi(M_1)=M_2$, $\Pi(P)=P$, and $\Pi^{-1}(A) = \Pi(A)$, for each point A of M.

It might be interesting to point out in connection with this definition that if PQ is a simple continuous arc, and M denotes the set $PQ-Q$, then $M-P$ may be expressed as the sum of two sets M_1 and M_2 and a continuous $(1-1)$ correspondence Π may be defined in such a way that all the properties of this correspondence that are necessary to make P fulfill the definition of center of symmetry of M, are fulfilled, excepting the last. This is illustrated by the following example.

Example 1. On the x-axis, let M denote the set of points $0 \leqq x < 1$. Let P be the origin. Let P_i $(i=1, 2, \cdots)$ be the point $1/2^i$, and let Q_i $(i=1, 2, \cdots)$ be the point $1-(1/2^{i+1})$. Let M_1 consist of the points P_{2i} and Q_{2i-1} and the segments Q_1P_1, $P_{2i}P_{2i+1}$, $Q_{2i+1}Q_{2i}$, for $i=1, 2, \cdots$. Finally let $M_2=M-P-M_1$. We shall define the continuous $(1-1)$ correspondence Π as follows: If $0 \leqq x \leqq 1/2$, then $\Pi(x)=x/2$; if $1/2 \leqq x \leqq 3/4$, then $\Pi(x)=x-(1/4)$; if $3/4 \leqq x < 1$, then $\Pi(x)=2x-1$. It is easily seen that $\Pi(M)=M$, $\Pi(M_1)=M_2$, $\Pi(P)=P$, but $\Pi^{-1}(A) \neq \Pi(A)$ for each point of $M-P$.

4. *Symmetrical Cut Points.* We shall say that the point P is a *symmetrical cut point of M,* if $M - P$ can be expressed as the sum of two mutually separated sets M_1 and M_2, which are such that there exists a continuous $(1-1)$ correspondence Φ having these properties: $\Phi(M_1 + P) = M_2 + P$ and $\Phi(P) = P$.

THEOREM 3. *If P is a symmetrical cut point of M, then P is a center of symmetry of M.*

Proof. If P is a symmetrical cut point of M, a continuous $(1-1)$ correspondence Π can be defined thus: for each point A of $M_1 + P$, $\Pi(A) = \Phi(A)$; for each point A of M_2, $\Pi(A) = \Phi^{-1}(A)$. The correspondence Π as defined thus has the properties that are necessary in order for P to be a center of symmetry of M.

THEOREM 4. *If M is a point set containing two points P and Q such that* (1) *$M - P$ is not connected and can be expressed as the sum of two mutually separated sets M_1 and M_2 which are homeomorphic, and* (2) *$M - Q$ is connected, then M is connected.*

Proof. Since $M - P$ is not connected and $M - Q$ is connected, the points P and Q are distinct. Hence the set $M - P = M_1 + M_2$ contains Q, and is accordingly not vacuous. But since $M_1 M_2 = 0$ and M_1 and M_2 are homeomorphic, it follows that $M - P$ contains at least two points. Hence $M - (P + Q)$ is not vacuous.

Suppose M_1 is the set containing Q, and let R be the point of M_2 which corresponds to Q under the correspondence between M_1 and M_2. If Q is not a limit point of $M - Q$, it is not a limit point of $M_1 - Q$, and hence R is not a limit point of $M_2 - R$. But in that case, the set $M - Q$ can be expressed as the sum of the two mutually separated sets $(M_1 - Q) + P + (M_2 - R)$ and R. But this is impossible, since by hypothesis $M - Q$ is connected.

Therefore Q is a limit point of the connected set $M - Q$, and the set M is connected.

COROLLARY 4A. *If P is a symmetrical cut point of M, and M contains a point Q such that $M - Q$ is connected, then M is connected.*

THEOREM 5. *If the point set M contains two distinct points A and B such that* (1) $M - A - B \neq 0$, (2) *every point of $M - A - B$ is a symmetrical cut point of M, and* (3) *$M - A$ is connected, then M is irreducibly connected between A and B.*

Proof. Let P be a point of $M - A - B$. Then by (2), P is a symmetrical cut point of M, and by (3) and Corollary 4A, M is connected.

Let M_1 be that one of the subsets of $M - P$ which contains A. If the set $M_1 + P - A$ were disconnected, it could be expressed as the sum of the two mutually separated sets H_1 and H_2, where H_2 contains P. But in that case $M - A = (M_1 + P + M_2) - A = (M_1 + P - A) + M_2$ could be expressed as the sum of the two mutually separated sets H_1 and $H_2 + M_2$, contrary to (3). Hence $M_1 + P - A$ is connected.

Hence $\Phi(M_1 + P - A) = M_2 + P - \Phi(A)$ is connected. Since $\Phi(A) \neq P$, it follows that $M - \Phi(A)$ is the sum of the two connected sets $M_1 + P$ and $M_2 + P - \Phi(A)$ having P in common, and $M - \Phi(A)$ is therefore connected. It follows from (2) that $\Phi(A) = B$. Hence $M - B$ is connected and B is a point of M_2.

In other words, each point P of $M - A - B$ separates A and B in M, and hence M is irreducibly connected between A and B.

5. *A New Characterization of a Simple Continuous Arc.* If M is a simple continuous arc from A to B, the hypotheses of Theorem 5 are obviously satisfied. The following example shows however, that a set may satisfy these hypotheses, without necessarily being an arc. We give in Theorem 6, certain additional conditions which are sufficient to characterize an arc.

Example 2. In a plane, let $A = (1, 0)$, $B = (-1, 0)$, $P_n = (n/n + 1, 0)$, $Q_n = (-n/n + 1,\ 0)$. Let E_n be the curve $y = \sin 1/(nx + x - n)(nx + 2x - n - 1)$, for $n/n. + 1 < x < n + 1/n + 2$, and let F_n be the curve $y = \sin 1/(nx + x + n)(nx + 2x + n + 1)$ for $-(n+1)/n + 2 < x < -n/n + 1$. Let $M = A + B + \sum_{i=0}^{i=\infty} (P_n + Q_n + E_n + F_n)$. It is easily seen that M is irreducibly connected between A and B, and that every point of $M - A - B$ is a symmetrical cut point of M.

THEOREM 6. *These three conditions are logically equivalent:*

(a) *M is a simple continuous arc from A to B;*

(b) *M is a closed point set containing two points A and B such that $M - A - B \neq 0$, every point of $M - A - B$ is a symmetrical cut point of M, and $M - A$ is connected;*

(c) *M is a regular (connected im kleinen) point set containing two points A and B such that $M - A - B \neq 0$, every point of $M - A - B$ is a symmetrical cut point of M, and $M - A$ is connected.*

The equivalence of conditions (a) and (b) follows from Theorem 5 and theorems due to Lennes and Hallett.* The equivalence of (a) and (c) follows from Theorem 5 and a theorem due to G. T. Whyburn.†

In comparing characterizations (b) and (c) with the previously known characterizations of an arc, it will be noted that boundedness is not assumed, nor is connectivity explicitly assumed. In many respects it resembles the characterization of Sierpinski,‡ with the homeomorphism of $M_1 + P$ and $M_2 + P$ taking the place of Sierpinski's condition of boundedness.

6. *Symmetrical Cut Points of an Unbounded Continuum.* Every point of an open curve M is a symmetrical cut point of M, but if it is known that every point of an unbounded continuum M is a symmetrical cut point of M, it does not follow that M is an open curve, as the following example will show.

Example 3.§ In a plane let K denote the positive half of the x-axis including $x = 0$. Let K_{10j} $(j = 0, \pm 1)$ denote the point set consisting of the segment from $(1, 0)$ to $(2, j)$ and all points of the line $y = j$ for $x \geqq 2$. Let K_{2ij} $(i = 0, \pm 1;\ j = 0, \pm 1)$ denote the set consisting of the segment from $(3, i)$ to $[4, (3i + j)/3]$ and all points of the line $y = (3i + j)/3$ for $x \geqq 4$. Continue this process. In general, let K_{nij} $[i = 0, \pm 1, \pm 2, \cdots, \pm(3^{n-1} - 1)/2;\ j = 0, \pm 1]$ denote the set consisting of the segment from $(2n - 1, i/3^{n-2})$ to $[2n, (3i + j)/3^{n-1}]$ and all points of the line $y = (3i + j)/3^{n-1}$ for $x \geqq 2n$. Let $N = K + \sum_{n=1}^{n=\infty} \sum_{i=R_n}^{i=+R_n} \sum_{j=-1}^{j=+1} K_{nij}$, where $R_n = (3^{n-1} - 1)/2$, and let N_i denote the set obtained by rotating N in the plane about the origin as center thru an angle of $\pi i/2$.

If $M = N_0 + N_1 + N_2 + N_3$, then M is an unbounded continuum, every point of which is a symmetrical cut point of M.

UNIVERSITY OF BUFFALO,
BUFFALO, NEW YORK.

---

* N. J. Lennes, *American Journal of Mathematics*, Vol. 33 (1911), p. 308; G. H. Hallett, Jr., *Bulletin of the American Mathematical Society*, Vol. 25 (1919), p. 325.

† *Bulletin of the American Mathematical Society*, Vol. 33 (1927), p. 688.

‡ *Annali di Matematica*, Ser. 3, Vol. 26 (1916), pp. 131-150.

§ The example is due to Professor J. R. Kline. In describing it, I have made use of the description of a more general set by W. L. Ayrès. See his paper in *Monatshefte für Mathematik und Physik*, Vol. 36 (1929), p. 136.

The Real Unit Segment as a Number Field.

By E. T. BELL.

In constructing an arithmetic of symmetric functions, it became necessary to devise an algebra, analysis and arithmetic of the numbers in the closed interval $[0, 1]$ isomorphic to the like, as commonly understood, for the closed interval $[-\infty, +\infty]$. As the solutions are of interest in themselves, one is given here independently of its applications.

1. Consider the real function $R(t)$,

$$R(t) = 1/2 + (1/\pi)\tan^{-1} t, \quad |\tan^{-1} t| \leqq \pi/2, \quad -\infty \leqq t \leqq +\infty.$$

Then the real function $R^{-1}(t)$ inverse to $R(t)$ is

$$R^{-1}(t) = -\cot \pi t, \qquad 0 \leqq t \leqq 1;$$

and

$$0 \leqq R(t) \leqq 1, \qquad -\infty \leqq R^{-1}(t) \leqq +\infty.$$

If a, b are any real numbers, $R(a+b)$, $R(a-b)$, $R(a \cdot b)$, $R(a \div b)$ are real numbers ($b \neq 0$ in the last).

We now define the operations $|+|$, $|\cdot|$ upon the numbers $R(t)$ by

$$R(a) \,|+|\, R(b) = R(a+b),$$
$$R(a) \,|\cdot|\, R(b) = R(a \cdot b).$$

Hence the inverse, $|-|$, of $|+|$, and the inverse, $|\div|$, of $|\cdot|$, are given by

$$R(a) \,|-|\, R(b) = R(a-b),$$
$$R(a) \,|\div|\, R(b) = R(a \div b),$$

in the last of which $b \neq 0$.

Write $R(a) = \alpha$, $R(b) = \beta$ for the moment. Then $0 \leqq \alpha \leqq 1$, $0 \leqq \beta \leqq 1$, $a = R^{-1}(\alpha)$, $b = R^{-1}(\beta)$. Hence, from the definition of $|+|$, it follows that

$$\alpha \,|+|\, \beta = R[R^{-1}(\alpha) + R^{-1}(\beta)],$$
$$= 1/2 - (1/\pi)\tan^{-1}[\cot \pi\alpha + \cot \pi\beta],*$$

and similarly for $|-|$, $|\cdot|$, $|\div|$.

The open interval derived from $[0, 1]$ by omitting both end points 0, 1, will be denoted by $\{0, 1\}$, and likewise for $[-\infty, +\infty]$, $\{-\infty, +\infty\}$.

* The principal value of any inverse tangent is always to be understood; thus $-\pi/2 < \tan^{-1}\phi \leqq +\pi/2$.

THEOREM 1. *The set of all numbers in* $\{0, 1\}$ *is a field, say* $\{R\}$, *with respect to* $|+|$, $|-|$, $|\cdot|$, $|\div|$ *(addition, subtraction, multiplication, division in* $\{R\}$*), defined by*

$$\begin{aligned} a \,|+|\, b &= 1/2 - (1/\pi)\tan^{-1}(\cot \pi a + \cot \pi b), \\ a \,|-|\, b &= 1/2 - (1/\pi)\tan^{-1}(\cot \pi a - \cot \pi b), \\ a \,|\cdot|\, b &= 1/2 + (1/\pi)\tan^{-1}(\cot \pi a \cot \pi b), \\ a \,|\div|\, b &= 1/2 + (1/\pi)\tan^{-1}(\cot \pi a \tan \pi b), \\ 0 < a < 1, &\qquad 0 < b < 1. \end{aligned}$$

The zero, $0'$, *and the unity,* $1'$, *of* $\{R\}$ *are*

$$0' = 1/2, \qquad 1' = 3/4.$$

The field $\{R\}$ *is isomorphic to the field of all numbers in* $\{-\infty, \infty\}$, *in which* $0, 1, +, -, \cdot, \div$ *have their usual meanings.*

THEOREM 2. *The ideal elements* $|+|\,\infty'$ *and* $|-|\,\infty'$ *(respectively the "positive infinity" and the "negative infinity" adjoined to* $\{R\}$*), having with respect to numbers and operations of* $\{R\}$ *the abstractly identical properties of* $+\infty$ *and* $-\infty$ *with respect to* $[-\infty, +\infty]$, *are*

$$|+|\,\infty' = 1, \qquad |-|\,\infty' = 0.$$

The result of adjoining $|+|\,\infty'$ and $|-|\,\infty'$ to $\{R\}$ will be denoted by $[R]$. The following makes Theorem 2 more explicit.

THEOREM 3. *The negative,* $|-|\,t$, *of any number* t *in* $[R]$ *is given by*

$$|-|\,t = 1 - t, \qquad 0 \leqq t \leqq 1;$$

the reciprocal, $1'\,|\div|\,t$, *is given by*

$$\begin{aligned} 1'\,|\div|\,t &= 1/2 - t, & \epsilon \leqq t \leqq 1/2 - \epsilon, \\ 1'\,|\div|\,t &= 3/2 - t, & 1/2 + \epsilon \leqq t \leqq 1 - \epsilon, \end{aligned}$$

where ϵ *is a positive real variable* $< 1/2$ *which attains the limit* 0.

2. Serial order is introduced into $[R]$ by the definitions of the relations $|>|$, $|<|$. We shall say that the number t in $[R]$ is *positive* in $[R]$, written $t\,|>|\,0'$, if and only if $t > 1/2$; if $t < 1/2$, we shall say that t is *negative* in $[R]$, and write $t\,|<|\,0'$. If a, b are in $[R]$, and if $a \neq b$, we shall say that *a is greater than b* in $[R]$, and write $a\,|>|\,b$, if $(a\,|-|\,b)\,|>|\,0'$. Similarly, *a is less than b* in $[R]$, $a\,|<|\,b$, if $(a\,|-|\,b)\,|<|\,0'$. *Equality* in R is indicated by $=$, instead of $|=|$, as the meaning is that of $=$ for real numbers.

From the definition of $a \,|\!-\!|\, b$ and the inequalities

$$\begin{array}{ll} \cot \pi a < \cot \pi b & \text{if } a > b, \\ \cot \pi a > \cot \pi b & \text{if } a < b, \\ 0 \leqq a \leqq 1, & 0 \leqq b \leqq 1, \end{array}$$

we have the

THEOREM 4. *If a, b are unequal in* $[R]$, $a \,|\!>\!|\, b$ *or* $a \,|\!<\!|\, b$ *according as* $a > b$ *or* $a < b$. *Hence to analysis for* $[-\infty, +\infty]$ *there is an isomorph for* $[0, 1]$ *in* $[R]$.

3. The theory of numbers in $[R]$ will follow from the definitions of integer and arithmetical divisibility in $[R]$.

By definition, $0'$ $(=1/2)$ is a rational integer (*the zeroth rational integer*) in $[R]$. From $0'$ we define the nth *positive integer* n' in $[R]$ inductively by n successive additions in $[R]$ of the unity $1'$ $(=3/4)$ in $[R]$; the nth *negative*, or the $-n$th, *integer* in $[R]$ is then $|\!-\!|\, n'$. Hence the mth integer m' in $[R]$ is

$$m' = R(m) \qquad (m \text{ an integer} \gtreqless 0).$$

Let c', b' be integers in $[R]$. Then c' is said to be *divisible by b' in* $[R]$ if and only if $c' \,|\!\div\!|\, b'$ is an integer in $[R]$, and we write $b' \mid c'$ when c' is divisible by b' in $[R]$. Hence $b' \mid c'$ when and only when c is divisible by b. If c', $|\!\pm\!|\, 1'$ are the only values of b' for which $b' \mid c'$, c' is called a *prime in* $[R]$ if $c' \neq 1'$. The positive primes in $[R]$ are therefore the $R(p)$, where p runs through all positive rational primes. Unique decomposition of integers in $[R]$ into primes in $[R]$ follows, and it is obvious that the whole theory of the rational integers goes over in simple isomorphism into a theory of rational integers in $[R]$. The extension to algebraic numbers and ideals in $[R]$ is immediate.

Concerning Collections of Continua not All Bounded.*

By J. H. Roberts.

In his paper *Concerning upper semi-continuous collections of continua* † R. L. Moore proved that if G is any upper semi-continuous collection of mutually exclusive bounded continua which fills the plane and no one of which separates the plane then G is a plane with respect to its elements. Moore suggested to me the study of what kinds of spaces would be obtained by removing the condition of the *boundedness* of the continua of G. In particular he raised this question: *If G is an upper semi-continuous collection of mutually exclusive continua which fills a plane S and no continuum of G separates S, then is G topologically equivalent to a subset of the plane?*

In the present paper an example is given which shows that if some continua of such an upper semi-continuous collection G are unbounded then the space of elements of G is not necessarily even *metric*. An example is also given of an upper semi-continuous collection G of continua filling the plane, no one of which separates the plane, such that G is topologically equivalent to a sphere. In this example only one continuum of the collection G is unbounded. It is then shown that with the additional hypotheses that the space of elements of G is metric and that G contains more than one unbounded continuum the answer to Moore's question is in the affirmative. Indeed the particular subsets of a plane which are topologically equivalent to such spaces of elements are completely characterized.

Definition 1.‡ If G is a collection of mutually exclusive continua then G is said to be an *upper semi-continuous* collection if for every sequence of

* Presented to the Society, Dec. 27, 1928.

† *Transactions of the American Mathematical Society*, Vol. 27 (1925), pp. 416-428.

‡ It is obvious that for the case where the continua of G are bounded, definitions 1 and 2 are equivalent to the following given by Moore (ibid.): "A collection G of continua is said to be an *upper semi-continuous* collection if for each element g of the collection G and each positive number e there exists a positive number d such that if x is any element of G at a lower distance from g less than d then the upper distance of x from g is less than e. The element p of such a collection G is said to be a *limit element* of the subcollection K of G if for every positive number e there exists some element of K which is distinct from p and whose upper distance from p is less than e."

continua $h_1, h_2, \cdots$ of G containing points $P_1, P_2, \cdots$ such that the sequence $P_1, P_2, \cdots$ has a sequential limit point P lying in a continuum g_P of G it is true that if $Q_1, Q_2, \cdots$ is a sequence of points such that for every i the point Q_i belongs to h_i then all the limit points of the point set $Q_1 + Q_2 + \cdots$ belong to the continuum g_P.

Definition 2. If G is an upper semi-continuous collection of continua then an element g of G is said to be a *limit element* of a set K of elements of G if and only if g contains a point P which is a limit point of the point set obtained by adding together all of the continua of K except g.

Example 1. Let r be a ray of an open curve. Let G be the collection of mutually exclusive continua filling the plane whose only non-degenerate * element is r. Clearly G is an upper semi-continuous collection. I will show that with respect to the above defined notion of limit element G is not a metric space. Let $C_1, C_2, C_3, \cdots$ be a set of concentric circles all containing points of the ray r, and such that for every n the diameter of C_n is n. For every n let L_n be an infinite sequence of points on C_n no one of which belongs to r, but such that some point of r is a sequential limit point of the sequence L_n.

Now suppose that G is metric. Then there exists a distance function $\delta(x, y)$ such that for every two distinct elements x and y of G we have $\delta(x, y) > 0$, and such that an element g of G is a limit element of a set K of elements if and only if for every positive number e the collection K contains an element k_e such that $\delta(k_e, g) < e$. Now for every n the element r is a limit element of the sum of the elements of the sequence L_n. Therefore for every n there exists an element p_n of L_n such that $\delta(r, p_n) < 1/n$. Then r is a limit element of the set of elements $p_1 + p_2 + p_3 + \cdots$. But no point of r is a limit point of the point set $p_1 + p_2 + p_3 + \cdots$. Hence the supposition that G is metric has led to a contradiction.

Example 2. (See fig. 1). Assume a rectangular coördinate system. If c is a real number ($c \geqq 2$) let g_c be the continuum consisting of the two intervals $x = \pm 1/c$, $0 \leqq y \leqq c$, together with the larger arc of the circle with the origin as center and radius equal to $[c^2 + 1/c^2]^{1/2}$ which has just its end points on these intervals. Let G be the collection of continua containing (a) the continuum g_c for every real number c ($c \geqq 2$), (b) the ray $x = 0$, $y \geqq 0$, and (c) every remaining *point* of the plane. It can be seen that G is an upper semi-continuous collection, and that it is metric with respect to

* An element is said to be *degenerate* if it contains but a single point. See R. L. Moore, "Concerning Upper Semi-Continuous Collections," *Monatshefte für Mathematik und Physik*, Vol. 36 (1929), pp. 81-88.

its elements. That G is not topologically equivalent to any subset of the plane follows from the fact that (by theorem III) G is topologically equivalent to a sphere.*

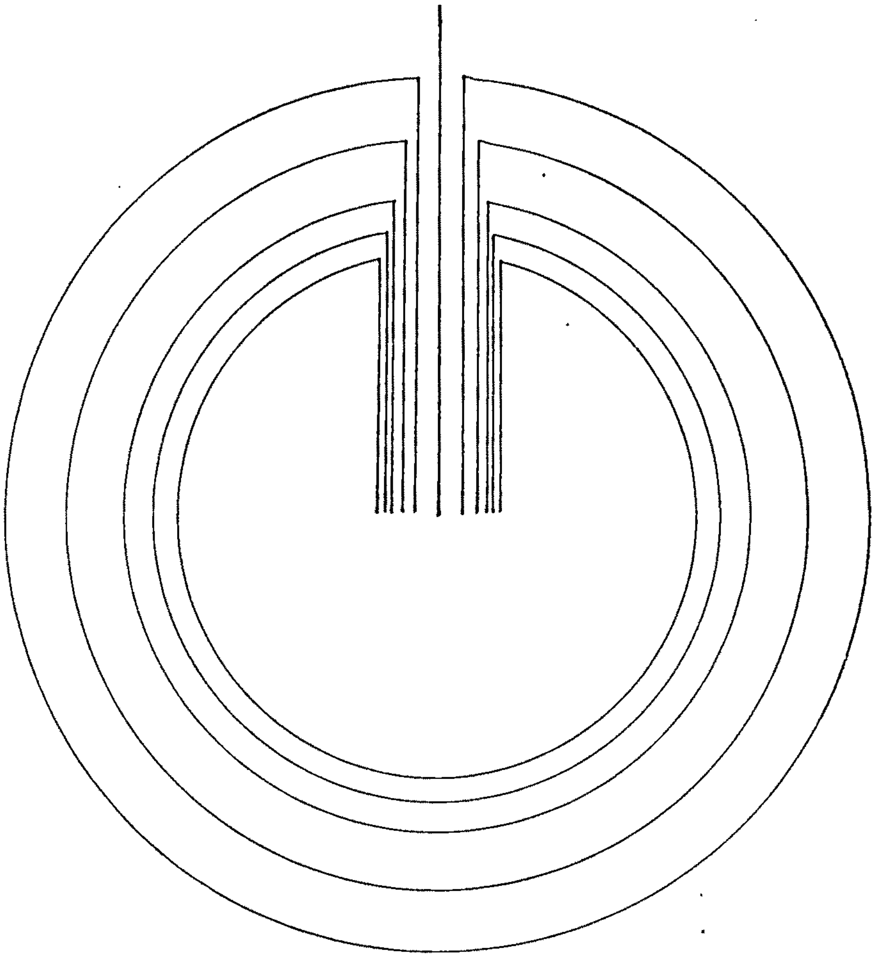


FIGURE 1.

THEOREM I.† *Let G be a collection of mutually exclusive unbounded*

* Another example is the following: Let G be a collection of mutually exclusive continua filling the plane of which the only non-degenerate element is a circle plus its exterior. Clearly G is topologically equivalent to a sphere. In the example given above no continuum of the collection G contains a domain.

† Dr. N. E. Rutt has obtained an independent proof of this theorem. In fact credit is due him for having been, as far as I know, the first to consider this proposition. Without knowing Dr. Rutt was working on this problem I obtained, as a proof of a somewhat different theorem, the argument herewith given.

continua whose sum is a closed point set M. If A and B are points and no continuum of G separates A from B then M does not separate A from B.

Proof. Suppose that the hypothesis of the theorem is satisfied but that M separates a point A from a point B. On the basis of this supposition I will show that M contains an *irreducible integral continuum* * which separates A from B.

In view of Mazurkiwicz's extension † of the Phragmen-Brouwer theorem to unbounded point sets it follows from Knaster and Kuratowski ‡ that every point set separating A from B contains a continuum separating A from B. In view of this and the additional fact that M is closed it is clear that some maximal connected subset of M is an *integral* continuum separating A from B. Let the interiors of the circles with rational radii and rational centers be ordered $R_1, R_2, \cdots$. For each positive integer n let H_n denote the sum of the continua of G which contain a point in R_n. Suppose that M does not contain an irreducible integral continuum separating A from B. Let n_1 be the smallest integer such that $M - H_{n_1}$ contains an integral continuum separating A from B, and let M_1 be such an integral continuum. Let n_2 be the smallest integer greater than n_1 such that $M_1 - M_1 \cdot H_{n_2}$ contains an integral subcontinuum separating A from B. It is clear that there exists a sequence of numbers $n_1, n_2, \cdots$ such that $n_1 < n_2 < n_3 < \cdots$ and a sequence of integral continua $M_1, M_2, M_3, \cdots$ such that for each k (a) the continuum M_k contains M_{k+1}, (b) M_k separates A from B and (c) M_k contains no point of $H_{n_1} + H_{n_2} + \cdots + H_{n_k}$ but if n is any integer less than n_k and is not any n_i $(i = 1, 2, \cdots, k-1)$ then every integral subcontinuum of M_{k-1} which separates A from B does contain a point of H_n. Now for every k every simple continuous arc from A to B contains a closed subset of M_k. Hence there exists at least one point common to all of the sets $M, M_1, M_2, \cdots$. Let Q denote the set of all points common to $M, M_1, M_2, \cdots$. Clearly Q is closed and is an integral subset of M. That Q separates A from B follows from the fact that Q contains a point on every arc from A to B. It follows from property (c) above that no integral continuum which is a proper subset of Q

* By an *integral continuum* is meant a continuum N such that if g is a continuum of G and N contains a single point of g then N contains every point of g. An integral continuum N is said to be an *irreducible integral continuum* (with respect to the property of separating A from B) if N separates A from B but no integral continuum of M which is a proper subset of N separates A from B.

† "Extension du théorème du Phragmen-Brouwer aux ensembles non bornés," *Fundamenta Mathematicae*, Vol. 3 (1922), p. 20.

‡ "Sur les ensembles connexes," *Fundamenta Mathematicae*, Vol. 2 (1921), p. 233.

separates A from B. But by Knaster and Kuratowski the set Q contains a continuum separating A from B. Hence Q *is* an irreducible integral continuum separating A from B.

Now there exist three continua g_1, g_2, and g_3 of G which belong to Q and three arcs AP_1, AP_2, and AP_3 such that A is the common part of each two of these three arcs, P_i belongs to g_i $(i=1, 2, 3)$ and no other point of the arc AP_i belongs to Q. Let D_1 denote that complementary domain of $\Sigma_1{}^3 AP_i + g_i$ which contains B. One continuum of the set g_1, g_2, g_3 contains no point in $\bar{D}_1$. Hence the point set $Q \cdot \bar{D}_1$ is a *proper* integral subset of Q and therefore does not separate A from B. Let BA denote an arc from B to A which has no point in common with $Q \cdot \bar{D}_1$. Let C denote the first point of the boundary of D_1 on BA in the order from B to A. Since the arc BA has no point in common with $Q \cdot \bar{D}_1$ the point C must belong to one of the arcs AP_1, AP_2, or AP_3, and must be distinct from P_1, P_2, and P_3. Obviously then there exists an arc from B to A which contains no point whatsoever of Q. Hence the supposition that M separates A from B has led to a contradiction.

COROLLARY 1. *Suppose P is a point, G is a collection of continua each two having in common just the point P, and M, the sum of the continua of G, is closed and bounded. Then if A and B are points and no continuum of G separates A from B the continuum M does not separate A from B.*

COROLLARY 2. *Suppose G is an upper semi-continuous collection of mutually exclusive continua filling the plane. Let T be the sum of all the unbounded continua of G. Then if A and B are points which are not separated by any unbounded continuum of G they are not separated by T.*

THEOREM II. *Let G be an upper semi-continuous collection of unbounded continua lying in a plane S and let M be a continuum of elements of G. Let g_1, g_2, g_3 and g_4 be any elements of M no two of which are separated in S by an element of G. Then the sum of two of the elements g_2, g_3, g_4, disconnect M between the third one and g_1.*

Proof. There exists a point A, four arcs AP_1, AP_2, AP_3 and AP_4 such that A is the common part of each two, P_i belongs to g_i $(i=1,\cdots,4)$ and no other point of AP_i belongs to any continuum g_j $(j=1,\cdots,4)$. Two of the arcs AP_i $(i=2, 3, 4)$ plus the corresponding continua of the set g_2, g_3, g_4 separate the third one from g_1. Suppose that Z $(Z=g_2+g_3$

$+ AP_2 + AP_3$) separates g_1 from g_4. (See fig. 2). Let D_1 and D_4 denote the complementary domains of Z containing g_1 and g_4, respectively. Let g denote any continuum of G in M distinct from g_2 and g_3. Not both of the sets $D_1 \cdot g$ and $D_4 \cdot g$ are unbounded. For since g does not separate g_2 from g_3 there exists a simple continuous arc P_2CP_3 with no point in g_2 or g_3 except P_2 and P_3, and no point in g. Let J be the interior of a circle which encloses the arcs P_2CP_3 and P_2AP_3. It can be shown that the continuum $g_2 + g_3 + P_2CP_3$ separates all points of D_1 and D_4 not in J. But g is connected and has no point in common with this continuum. Hence g has points in only one of the sets $D_1 \cdot (S - \bar{J})$ and $D_4 \cdot (S - \bar{J})$.

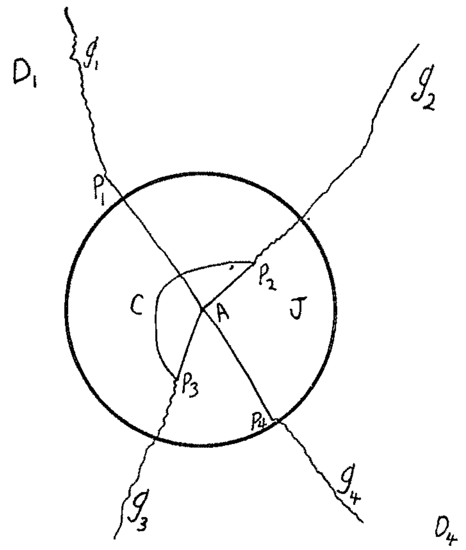


FIGURE 2.

Let S_1 denote the point set consisting of the sum of all continua of G in M which are unbounded in the domain D_1, and S_2 the sum of all those unbounded in D_4. The set S_1 contains g_1 and the set S_2 contains g_4. Suppose that a point E of S_1 is a limit point of S_2. Then there exists a sequence of continua $h_1, h_2, \cdots$ lying in S_2 containing a sequence of points $E_1, E_2, \cdots$ which has E as sequential limit point. If g_E denotes the continuum of G containing E, in view of the fact that G is upper semi-continuous, it is clear that $g_E \cdot D_4$ is unbounded. But g_E is in S_1. This contradiction shows that no point of S_1 is a limit point of S_2. Similarly no point of S_2 is a limit point of S_1. Hence g_1 and g_4 are separated in M by the sum of the continua g_2 and g_3.

LEMMA 1. *Suppose G is an upper semi-continuous collection of continua lying in a plane S and G is metric. Then if $\delta(x,y)$ denotes a distance function, for each element g of G there exists a circle C_g and a positive number e_g such that if x is an element of G for which $\delta(x,g) < e_g$ then x contains a point inside the circle C_g.*

Proof. The function $\delta(x,y)$ satisfies the following conditions: (1) $\delta(x,y) > 0$ if $x \neq y$, (2) $\delta(x,x) = 0$, (3) $\delta(x,y) + \delta(y,z) \geqq \delta(x,z)$, and (4) an element g of G is a limit element of a set K if and only if for every positive number e there exists an element g_e of K for which $\delta(g, g_e) < e$. Let P be a point of the plane, and for each positive integer n let C_n denote the circle with P as center and n as radius. Let g denote any element of G. Suppose that for every n there exists an element x_n of G such that $\delta(x_n, g) < 1/n$ and x_n has no point within the circle C_n. Then the limit of $\delta(x_n, g)$, as n increases indefinitely, is zero, whence g is a sequential limit element of the sequence $x_1, x_2, \cdots$. But this is impossible since no sequence of points $P_1, P_2, \cdots$ where for every n the point P_n belongs to x_n, has a sequential limit point. This contradiction proves the lemma.

THEOREM III. *If G is an upper semi-continuous collection of continua filling the plane S such that no continuum of G separates S and G is metric, then in order that G be topologically equivalent to a sphere it is necessary and sufficient that exactly one continuum of G be unbounded.*

Proof. The necessity of the condition follows from these two facts: (1) If *no* continuum of G is unbounded then G is * equivalent to the whole plane S. (2) If at least two continua of G are unbounded then G is † equivalent to a proper subset of the plane.

The condition is sufficient. Let g be the continuum of G which is unbounded. The point set $S - g$ is a simply connected domain and is therefore topologically equivalent to the plane S. Hence by Moore's theorem (ibid.) the collection $G - g$ ‡ is topologically equivalent to a plane and therefore to a sphere minus one point.

By lemma 1 there exists a circle C and a positive number e such that if x is an element of G and $\delta(x, g) < e$ then x has a point inside the circle C. Since G is upper semi-continuous and g is the only unbounded continuum of G it follows that if C_1 is any closed subset of C which does not contain

* R. L. Moore, *loc. cit.*

† This follows from theorem IV, the proof of which does not depend upon theorem III.

‡ That is, the set of all continua of G except g.

a point of g, and K is the sum of the elements of G which contain points of C_1, then K is a bounded point set. It therefore follows that every unbounded set of elements of $G - g$ has g as limit element. Hence G is topologically equivalent to a sphere.

LEMMA 2. *Let G be an upper semi-continuous collection of continua filling the plane, none separating the plane, such that G is metric. Then if T is the sum of all unbounded continua of G and M is a maximal connected subset of T it follows that no point of M is a limit point of $T - M$.*

Proof. Suppose that lemma 2 is not true. Then there exists an element g and a sequence of elements $g_1, g_2, \cdots$ with g as sequential limit element such that all of the continua $g, g_1, g_2, \cdots$ are unbounded and no two of them belong to the same component of T. By lemma 1 there exists a circle C and a positive number e such that if x is a continuum of G and $\delta(x, g) < e$ then x contains a point inside the circle C. With the help of the methods used in the proof of theorem II it can be shown that there exists a subsequence $g_{n_1}, g_{n_2}, \cdots$ of the sequence $g_1, g_2, \cdots$ and for each k an arc A_kB_k of the circle C having just A_k in g and B_k in g_{n_k}, and a domain D_k which is one of the complementary domains of $g + g_{n_k} + A_kB_k$, such that, T_k denoting the sum of all continua of G which are unbounded in $\bar{D}_k$, the set T_k contains the set T_{k+1}. It can also be shown that if $\bar{e}$ is any positive number then there exists an integer m such that if x is an unbounded continuum of G which contains a point of C and $\delta(x, y) > \bar{e}$ then x does not belong to T_m.

Let m be an integer such that if x is a continuum of G which is a subset of T_m, and x contains a point of C, then $\delta(x, g) < e/2$. Since T_m contains continua of the sequence $g_1, g_2, \cdots$ and also contains g it follows that it is not connected. Therefore there exists an unbounded continuum r of G such that $T_m = H_g + H_r$, where H_g and H_r are mutually exclusive closed sets containing g and r, respectively. (See fig. 3). Since H_g and H_r are closed, and G is an upper semi-continuous collection, there exist two open sets R_g and R_r such that (1) $\bar{R}_g \cdot \bar{R}_r = 0$, (2) R_g contains $C \cdot H_g$ and R_r contains $C \cdot H_r$, and (3) if x is any continuum of G with a point in R_g [R_r] then there exists a continuum k of G which belongs to H_g [H_r] and has a point on C such that $\delta(x, k) < e/2$. Let P denote the center of the circle C and for each positive integer s let a_s be an arc lying in D_m except that its end points belong to g and r, respectively, and such that every point of a_s is at a distance greater than s from P. Since every continuum of the upper semi-continuous collection G which contains a point of C not in R_g or R_r is bounded in the domain D_m and since moreover the sum of all such continua is closed

(R_g and R_r being open) it follows that there exists an integer t such that if $s > t$ then no continuum of G which contains a point of a_s can contain a point of C not belonging to $R_g + R_r$.

For every integer s greater than t let Q_s denote the first point of a_s in the order from r to g which belongs to a continuum of G which contains a point of $\bar{R}_g$. Let x_s denote that continuum of G which contains Q_s. Since $\delta(x_s, g) < e$, there exists a point P_s of a_s which preceeds Q_s in the order from r to g, such that if y_s is any continuum of G which contains a point of the

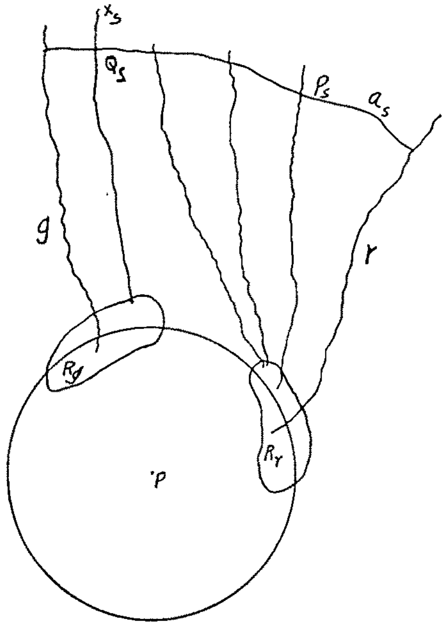


FIGURE 3.

arc P_sQ_s, then $\delta(x_s, g) < e$, whence y_s contains a point of $\bar{R}_g + \bar{R}_r$. But if y_s is different from x_s it does not contain a point in $\bar{R}_g$. Therefore there exists a sequence of continua of G, $h_1, h_2, \cdots$ with x_s as sequential limit element and such that for every integer k the continuum h_k contains a point in $\bar{R}_r$. Therefore, since G is upper semi-continuous, x_s contains a point of both $\bar{R}_g$ and $\bar{R}_r$. Since for every s $(s > t)$ the continuum x_s has a point in the closed and bounded set $\bar{R}_g$ it follows that there exists a continuum z of G which is a limit element of the sum of the elements of the sequence x_{t+1}, $x_{t+2}, \cdots$. Obviously z is unbounded in $\bar{D}_m$. Moreover it contains a point both of $\bar{R}_g$ and of $\bar{R}_r$. But this is contrary to the fact that T_m, the sum of all continua of G which are unbounded in $\bar{D}_m$, is the sum of the mutually separated sets H_g and H_r, where H_g contains no point of $\bar{R}_r$ and H_r contains

no point of $\bar{R}_g$. Hence our supposition that lemma 2 is false has led to a contradiction.

LEMMA 3. *If G is an upper semi-continuous collection of continua which fills the plane, no continuum of G separates the plane but at least two continua of G are unbounded, and G is metric, then no element of G is a component of the sum of all the unbounded continua of G.*

Proof. Suppose that g is an unbounded continuum of G. Let k be any other unbounded continuum of G and let $\bar{e}$ be $\delta(g, k)$. Let e be any positive number less than $\bar{e}$. By lemma 1 there exists a circle C and a positive number e_1 ($e_1 < e$) such that if x is a continuum of G and $\delta(g, x) < e_1$ then x contains a point inside the circle C. Let P be the center of C and for every positive integer n let a_n be an arc with end points on g and k, respectively, every point of which is at a distance greater than n from P. Since $\delta(g, k) > e_1$ it follows that there exists a continuum x_n of G containing a point of a_n and such that $e_1/2 < \delta(x_n, g) < e_1$. For every n the continuum x_n has a point within C and a point at a distance greater than n from P. Hence if for each n the continuum x_n is bounded the sum of the elements of the sequence $x_1, x_2, \cdots$ has a limit element x_e, which is unbounded. Since $e_1/2 < \delta(x_n, g) < e_1$ it follows that $\delta(x_e, g) \leqq e_1 < e$, and x_e is distinct from g. Hence for every positive number e there exists an unbounded continuum x_e of G distinct from g such that $\delta(x_e, g) < e$. Lemma 3 therefore follows in view of lemma 2.

LEMMA 4. *Under the hypotheses of lemma 3 if M is a collection of continua of G whose sum is a component of the sum of all unbounded continua of G then M is a simple closed curve or an open curve.*

Proof. It readily follows from theorem II that M is an atriodic* continuous curve. Hence M is an open curve, a simple closed curve, an arc, or a ray of an open curve. Let g denote any element of M. Let g_1, g_2 and g_3 denote three other elements of M such that the sum of g and g_2 separates g_1 from g_3 in M. Let AB denote an arc having only A in common with g and only B in common with g_2. Let D be either of the two complementary domains of $g + g_2 + AB$. By a slight modification of the proof of lemma 3

* If O, A_1, A_2 and A_3 are distinct points and OA_1, OA_2 and OA_3 are arcs each two having only O in common, then the continuum $OA_1 + OA_2 + OA_3$ is called a *triod*. A continuous curve which does not contain a triod is said to be *atriodic*. See R. L. Moore, "Concerning Triods in the Plane and the Junction Points of Plane Continua," *Proceedings of the National Academy of Sciences*, Vol. 14 (1928), pp. 85-88.

t can be seen that g is a limit element of the elements of M which are unounded in D. Therefore g is not an end element of M. Hence M contains o end elements and is therefore a simple closed curve or an open curve.

THEOREM IV. *If G is an upper semi-continuous collection of continua hich fills a plane S and is metric, at least two continua of G are unbounded nd no one of them separates S, then G is topologically equivalent to a domain) plus its boundary where* (a) *D is the interior of a circle, or* (b) *D is an nbounded domain whose boundary consists of one or more open curves no ne containing a limit point of the sum of the others.*

Proof. Case 1. Suppose first that some collection M of unbounded coninua of G is a simple closed curve. Suppose g is any element of M and h is any other unbounded continuum of G. Since g is not an end element f M it follows with the help of the methods used to prove theorem II that $g + h$ separates $M + h$. Since no element of M separates M it follows that h belongs to M; that is, every unbounded continuum of G is an element of M. Let P be any point and for every positive integer n let C_n be the circle with P as center and radius n. From lemma 1 it follows that for any element g of M there exists a positive integer k_g such that if x is a continuum of G nd $\delta(x, g) < 1/k_g$, then x contains a point inside the circle C_{k_g}. Since moreover every infinite set of elements of M has a limit element (M being a simple losed curve) it follows that there exists a single positive integer k such that f g is any element of M and x is a continuum of G such that $\delta(x, g) < 1/k$, hen x contains a point inside the circle C_k.

Let $g_1, g_2, \cdots$ denote any infinite sequence of distinct elements of G. f for any n all but a finite number of these continua contain a point inside C_n then the set of elements $g_1 + g_2 + \cdots$ has a limit element. Otherwise or every n $(n > k)$ let r_n be an integer such that g_{r_n} has no point inside C_n. Obviously $\delta(g_{r_n}, g) \geqq 1/k$ for every element g of M. It follows that there xists a continuum h_n of G such that $1/2k < \delta(h_n, g)$ for every element g of M, but $\delta(g, h_n) < 1/k$ for some element g of M, and h_n contains a point utside C_n. The sequence $h_{k+1}, h_{k+2}, \cdots$ obviously contains a subsequence with a sequential limit element h. Clearly h is unbounded. But since $1/2k < \delta(h_n, g)$ for every unbounded continuum g of G we have $1/2k \leqq \delta(h, h)$. This contradiction shows that for some integer n every element of G has a point inside the circle C_n so that G is compact.

Now every element of M is a limit element of the set of elements $G - M$. The sum of all the continua of $G - M$ is a connected domain D with a connected boundary. Hence D is topologically equivalent to the whole plane

S, and by Moore's theorem referred to above the collection $G - M$ is topologically equivalent to the plane S, or to the interior of a circle. Its boundary is equivalent to a circle. This completes the proof for case 1.

Case 2. Suppose that no collection of unbounded continua of G is a simple closed curve. Let T be the sum of all unbounded continua of G and let $M_1, M_2, \cdots$ be the subcollections of continua of G, the sums of whose continua form the components of T. In view of theorem I it can be seen that $S - T$ is a simply connected domain D. Hence in view of Moore's theorem (ibid.) the collection $G - (M_1 + M_2 + \cdots)$ is a simply connected domain. Since in addition for every i every element of M_i is a limit element of $G - M_i$ theorem IV is established.

THEOREM V. *If K is a point set described in the conclusion of theorem IV then there exists an upper semi-continuous collection G of continua which fills the plane such that* (1) *no continuum of G separates the plane and* (2) *with respect to its elements G is topologically equivalent to the point set K.*

Proof. For the case where K is a circle C plus its interior let the elements of G be the *points* interior to C together with the subsets of all rays starting at the center of C which lie on and outside of C. Obviously G is metric, and is topologically equivalent to C plus its interior. Suppose K is a domain D plus its boundary, where D is a connected domain bounded by open curves $N_1, N_2, \cdots$, such that for each i no point of N_i is a limit point of the sum of the other open curves of the sequence $N_1, N_2, \cdots$. For each i let D_i be the complementary domain of N_i which does not contain D. It can be shown that there exists an upper semi-continuous set G_i of rays of open curves lying in $\bar{D}_i$, each starting on N_i, and such that every point of $\bar{D}_i$ belongs to some ray of the set G_i. Let G be the set of all points of D together with all continua of the collection G_i for every i. The set G_i is topologically equivalent to the open curve N_i. It therefore follows that the collection G satisfies the conclusion of the theorem.

THE UNIVERSITY OF TEXAS.

Limited Tri-Linear Forms in Hilbert Space.*

By William L. Hart.

1. *Introduction.* At the foundation of Hilbert's theory of functions of infinitely many variables we find the notion of a limited † function, and, in particular, the idea of a limited bi-linear form in infinitely many variables. Similarly, one may consider limited tri-linear forms, or, in general, limited m-linear forms where m is a positive integer. The most fundamental properties of limited m-linear forms are easily anticipated by analogy with bi-linear forms. For a given $m > 2$, the proofs of these properties can in general be developed along the same lines as the existing demonstrations of results for bi-linear forms.‡ In this connection, however, one important exception is to be noted; the vital theorem on the uniform finiteness of a bi-linear form, which converges over the unit sphere in Hilbert space, has been established § by a method which demands essential modification in order to apply to the case of an m-linear form with $m > 2$.

The main purpose of this paper is to demonstrate the theorem on uniform finiteness for tri-linear forms by use of a method which applies as well in the consideration of m-linear forms, for any $m \geqq 2$. It is emphasized that this method, which is presented in Section 3, has merit solely because of its applicability when $m > 2$; its use when $m = 2$ would lead to much more complication than is found in the classical proof for this case. As a minor part of the paper, in Section 2 there is given a summary of other properties of tri-linear forms with details of the proofs only in a few instances where there is novelty as compared with the bi-linear case. In the entire discussion, all number symbols are supposed to be real valued.

2. *Limited Tri-linear Forms.* In the future, any small greek letter will denote a vector with infinitely many coördinates. Thus $\xi = (x_1, x_2, \cdots)$. By the n-th *section* of a vector ξ we shall mean the vector $\xi^{(n)} = (x_1, x_2, \cdots, x_n, 0, 0, \cdots)$, where all coördinates beyond the n-th one are zeros. If ξ is in

* Presented to the American Mathematical Society, June 20, 1929.

† David Hilbert, *Grundzüge Einer Allgemeinen Theorie der Linearen Integralgleichungen* (B. G. Teubner), p. IV.

‡ For definitions and theorems relating to limited linear and bi-linear forms, see Hellinger-Toeplitz, *Mathematische Annalen*, Vol. 69 (1910), p. 289.

§ Hellinger-Toeplitz, *loc. cit.*, p. 321; the details at the top of p. 324 do not generalize to the case of m-linear forms, with $m > 2$.

Hilbert space, the positive square root of the convergent series $\sum_{i=1}^{\infty} x_i^2$ will be denoted by $M(\xi)$ and will be called the modulus of ξ. We shall let H represent Hilbert space, and H_1 the set of vectors ξ where $M(\xi) \leqq 1$.

In accordance with Hilbert's definition of a limited function, we shall say that a three dimensional array $A \equiv (a_{ijk})$, and also the corresponding purely formal tri-linear expression

$$(1) \qquad A(\xi, \rho, \sigma) = \sum_{i,j,k=1}^{\infty} a_{ijk} x_i r_j s_k,$$

is *limited,* in case there exists a bound $E > 0$ such that for all ξ, ρ, and σ in H, and for all integers h, m, and n,

$$(2) \qquad | A(\xi^{(h)}, \rho^{(m)}, \sigma^{(n)}) | \leqq EM(\xi)M(\rho)M(\sigma).$$

An equivalent definition would result if (2) were altered by letting $h = m = n$. For abbreviation, we shall say that A is E-limited if (2) holds.

If any subscript of a_{ijk} is held fast, while the other two vary independently from 1 to ∞, we obtain a component matrix of A. If two subscripts are held fast, while the third varies from 1 to ∞, we obtain a component vector. The following Property I is easily proved by use of (2) and the corresponding definitions of limited linear and bi-linear forms; Properties II to IV can be demonstrated by direct methods like those used in similar connections in the consideration of bi-linear forms.

Property I. If A is E-limited, then each component vector, and each component matrix of A is E-limited.

Property II. If A is E-limited, then, if ρ and σ are in H, $\sum_{j,k=1}^{\infty} a_{ijk} r_j s_k$ *converges uniformly with respect to the index i.*

Property III. If A is E-limited, then, for all ρ and σ in H, the vector $\left(\sum_{j,k=1}^{\infty} a_{ijk} r_j s_k\right)$ *lies in H and its modulus is at most $EM(\rho)M(\sigma)$.*

Property IV. If A is E-limited, then, for every σ in H, the matrix $\left(\sum_{k=1}^{\infty} a_{ijk} s_k\right)$ *is limited, with $EM(\sigma)$ as a bound.*

We can consider evaluating a tri-linear form $A(\xi, \rho, \sigma)$ in 13 different ways, either as a triple series, or as an iterated series in twelve different ways.

Property V. If any one of the twelve iterated series for $A(\xi, \rho, \sigma)$ converges for all (ξ, ρ, σ) in H, then all of these twelve series converge and have the same sum for all (ξ, ρ, σ) in H.

Recognize that in the statement of Property V it is not assumed that A is limited. Let us discuss only the special case where we suppose that

$$(3) \qquad W(\xi, \rho, \sigma) \equiv \sum_{i=1}^{\infty} x_i \sum_{j,k=1}^{\infty} a_{ijk} r_j s_k$$

converges for all (ξ, ρ, σ) in H. This implies that $\sum_{j,k}$, on the right in (3), converges for all (ρ, σ) in H. Therefore, by virtue of the property of uniform finiteness for a bi-linear form, it follows that, supposing i to be fixed, the matrix (a_{ijk}) is limited. Hence, the double sum $\sum_{j,k}$ can be changed to either possible iterated form. Thus, $A(\xi, \rho, \sigma)$ converges to $W(\xi, \rho, \sigma)$ under each of the following methods of iterated summation: $\sum_i[\sum_j(\sum_k)]$; $\sum_i[\sum_k(\sum_j)]$. Since the last method converges for all (ξ, σ) in H when ρ is held fast, it follows from the uniform finiteness of a bi-linear form that the matrix $(\sum_{j=1}^{\infty} a_{ijk} r_j)$ is limited. Then, starting from $\sum_i[\sum_k(\sum_j)]$, the fundamental convergence properties of a limited bi-linear form show that $A(\xi, \rho, \sigma)$ converges to $W(\xi, \rho, \sigma)$ under each of the following methods of iteration: $\sum_{i,k}(\sum_j)$; $\sum_k[\sum_i(\sum_j)]$. In this fashion, we show that all of the 12 iterated forms for $A(\xi, \rho, \sigma)$ converge to $W(\xi, \rho, \sigma)$.

Property VI. If A is E-limited, then, for all (ξ, ρ, σ) in H, the formal series for $A(\xi, \rho, \sigma)$ converges to an unique value when evaluated in any of the thirteen possible ways, and

$$(4) \qquad | A(\xi, \rho, \sigma) | \leqq EM(\xi)M(\rho)M(\sigma).$$

The methods used in establishing the convergence properties of a limited bi-linear form would easily demonstrate that, if A is E-limited, then the particular iterated series $W(\xi, \rho, \sigma)$ of (3) converges and

$$(5) \qquad | W(\xi, \rho, \sigma) | \leqq EM(\xi)M(\rho)M(\sigma).$$

Hence, it follows from Property V that (4) is true if $A(\xi, \rho, \sigma)$ is evaluated as an iterated series in any of the 12 possible ways. To complete the proof of Property VI, it would be necessary to show that $A(\xi, \rho, \sigma)$ converges to $W(\xi, \rho, \sigma)$ when evaluated as a triple series; that is, it should be proved that

$$(6) \qquad \lim_{n \to \infty} \sum_{i,j,k=1}^{n} a_{ijk} x_i r_j s_k = \sum_{i=1}^{\infty} x_i \sum_{j,k=1}^{\infty} a_{ijk} r_j s_k,$$

where the right member, $W(\xi, \rho, \sigma)$, is known to converge. We omit the proof of (6), which is easily accomplished by the aid of (5).

Three successive applications of the Schwarz inequality * lead to the

* Hellinger-Toeplitz, *loc. cit.*, p. 293.

following result, which corresponds to a well known property of bi-linear forms.

Property VII. If the expression $V \equiv \sum_{i,j,k=1}^{\infty} a^2_{ijk}$ *converges, then A is limited, with* $V^{1/2}$ *as a bound.*

By reference to the definitions of continuity † which apply to functions defined in Hilbert space, and by use of the preceding properties of tri-linear forms, standard methods suffice to establish the following facts:

Property VIII. If $A(\xi, \rho, \sigma)$ *is limited, then it is continuous simultaneously in its arguments for all points in the region S defined by the following inequalities, where* g_1, g_2, *and* g_3 *are any assigned positive numbers:*

$$M(\xi) \leqq g_1; \qquad M(\rho) \leqq g_2; \qquad M(\sigma) \leqq g_3. \tag{7}$$

If $\sum_{i,j,k=1}^{\infty} a^2_{ijk}$ *converges, then* $A(\xi, \rho, \sigma)$ *is completely continuous in its arguments for all points* (ξ, ρ, σ) *satisfying* (7).

We shall call $C(\xi) \equiv \sum_{i,j,k=1}^{\infty} c_{ijk} x_i x_j x_k$ a cubic form when c_{ijk} is a symmetric function of (i, j, k). A cubic form is said to be limited, with E as a bound, in case

$$| C(\xi^{(n)}) | \leqq EM^3(\xi), \tag{8}$$

for every ξ in H and for every n.

Property IX. A necessary and sufficient condition that $C(\xi)$ *be limited is that the corresponding tri-linear form* $A(\xi, \rho, \sigma) \equiv \sum_{i,j,k=1}^{\infty} c_{ijk} x_i r_j s_k$ *be limited.*

It is obvious that, if $A(\xi, \rho, \sigma)$ is limited, then $C(\xi)$ is limited. To prove the converse, we note that the following identity is true:

$$\begin{aligned} 24A(\xi^{(n)}, \rho^{(n)}, \sigma^{(n)}) \equiv {} & [C(\xi^{(n)} + \rho^{(n)} + \sigma^{(n)}) - C(\xi^{(n)} - \rho^{(n)} + \sigma^{(n)}) \\ & - C(\xi^{(n)} + \rho^{(n)} - \sigma^{(n)}) + C(\xi^{(n)} - \rho^{(n)} - \sigma^{(n)})]. \end{aligned} \tag{9}$$

This identity can be verified by expanding the finite triple sums occurring, combining terms, and recalling that c_{ijk} is a symmetric function of (i, j, k). From (9) we see that, for all (ξ, ρ, σ) whose moduli are at most 1,

$$| A(\xi^{(n)}, \rho^{(n)}, \sigma^{(n)}) | \leqq 108E/24, \tag{10}$$

† Cf. Hellinger-Toeplitz, *loc. cit.*, p. 307-308.

if C is E-limited. Consequently, A is limited, with $36E/8$ as a bound.

3. *Uniform Finiteness of a Tri-Linear Form.*

Property X. If any one of the thirteen methods for evaluating a tri-linear form $A(\xi, \rho, \sigma)$ converges for all (ξ, ρ, σ) in H, then A is limited.

The case of Property X where we assume that $A(\xi, \rho, \sigma)$ converges as a triple series, for all (ξ, ρ, σ) in H, can be treated by essentially the same simple process which serves to prove the corresponding result for a bi-linear form, which is presumed to converge as a *double* series. Presuming that the instance of the triple series has been treated, it follows from Property V that we can complete the proof of Property X by discussing merely the case where the special iterated series $W(\xi, \rho, \sigma)$ of (3) is assumed to converge for all (ξ, ρ, σ) in H. Under this hypothesis, to show that A is limited, it will be sufficient to prove that $|W(\xi, \rho, \sigma)|$ is bounded in H_1, because

$$(11) \qquad \Big| \sum_{i,j,k=1}^{n} a_{ijk} x_i r_j s_k \Big| = | W(\xi^{(n)}, \rho^{(n)}, \sigma^{(n)}) | .$$

To establish the desired result by an indirect process, let us suppose that $W(\xi, \rho, \sigma)$ is *not* bounded in H_1; it follows that there exists an infinite sequence of points $(\theta_p, \phi_p, \psi_p)$ in H_1 such that $\lim_{p\to\infty} | W(\theta_p, \phi_p, \psi_p) | = +\infty$. Let us agree to change the signs of all coördinates of those vectors θ_p for which $W(\theta_p, \phi_p, \psi_p) < 0$; then, $W(\theta_p, \phi_p, \psi_p) \geqq 0$ for all p and

$$(12) \qquad \lim_{p\to\infty} W(\theta_p, \phi_p, \psi_p) = +\infty .$$

We shall proceed to a contradiction by defining vectors (η, ζ, τ) in H_1 for which $W(\eta, \zeta, \tau)$ does not converge.*

For future use, recall that, in the discussion of Property V, it was seen that, if $W(\xi, \rho, \sigma)$ converges for all (ξ, ρ, σ) in H, then each component matrix of A is limited.

LEMMA 1. *If m and n are assigned integers > 0, there exists a $d > 0$, where $d \leqq 1$, such that, if (ξ, ρ, σ) are points in H_1 for which $M[\rho - \rho^{(m)}] \leqq d$ and $M[\sigma - \sigma^{(m)}] \leqq d$, then*

$$(13) \qquad | W(\xi^{(n)}, \rho, \sigma) - W(\xi^{(n)}, \rho^{(m)}, \sigma^{(m)}) | \leqq 1.$$

On applying the Schwarz inequality with respect to i in $W(\xi^{(n)}, \rho, \sigma)$ as given by (3), we obtain

* The same notion is the basis of the proof given by Hellinger-Toeplitz, *loc. cit.*, p. 321, in establishing the uniform finiteness of a bi-linear form.

$$(14) \qquad | W(\xi^{(n)}, \rho, \sigma) |^2 \leqq (\sum_{i=1}^{n} x_i^2) \sum_{i=1}^{n} (\sum_{j,k=1}^{\infty} a_{ijk} r_j s_k)^2.$$

Let K be a common bound for the n limited matrices $A_1, \cdots, A_n$ where $A_i \equiv (a_{ijk})_{j,k=1,2,\ldots}$. Then, if ξ is in H_1, from (14) we obtain

$$(15) \qquad | W(\xi^{(n)}, \rho, \sigma) | \leqq K n^{1/2} M(\rho) M(\sigma).$$

We employ (15) twice in the following reduction:

$$\begin{aligned} | W(\xi^{(n)}, \rho, \sigma) - W(\xi^{(n)}, \rho^{(m)}, \sigma^{(m)}) | \\ \leqq | W(\xi^{(n)}, \rho - \rho^{(m)}, \sigma) | + | W(\xi^{(n)}, \rho^{(m)}, \sigma - \sigma^{(m)}) | \\ \leqq K n^{1/2} [M(\sigma) M(\rho - \rho^{(m)}) + M(\rho^{(m)}) M(\sigma - \sigma^{(m)})]. \end{aligned}$$

Hence, the statement of Lemma 1 is satisfied if d is chosen as the minimum of 1 and $1/2Kn^{1/2}$.

To proceed with the construction of (η, ζ, τ), select a sequence (c_{1j}) where $c_{1j} > 0$ and $\sum_{j=1}^{\infty} c^2_{1j} \leqq 1$. On account of (12), there exist vectors $(\alpha_1, \beta_1, \gamma_1)$ in H_1 such that $W(\alpha_1, \beta_1, \gamma_1) > 2/c^3_{11}$. Let $\xi_1 = c_{11}\alpha_1$, $\rho_1 = c_{11}\beta_1$, and $\sigma_1 = c_{11}\gamma_1$: then, $(\xi_1, \rho_1, \sigma_1)$ are in H_1 and $W(\xi_1, \rho_1, \sigma_1) > 2$. Therefore, there exist integers $m_1 > 0$ and $n_1 > 0$ such that $W(\xi_1{}^{(n_1)}, \rho_1{}^{(m_1)}, \sigma_1{}^{(m_1)}) > 2$. Now, let us define a block of the coördinates (y_i, z_j, t_j) of (η, ζ, τ) as follows, where (x_{i1}, r_{j1}, s_{j1}) are, respectively, coördinates of $(\xi_1, \rho_1, \sigma_1)$:

$$(16) \qquad y_i = x_{i1}, \ (i = 1, 2, \cdots, n_1); \ z_j = r_{j1} \text{ and } t_j = s_{j1}, \ (j = 1, 2, \cdots, m_1).$$

We note that $M(\eta^{(n_1)}) \leqq c_{11} \leqq 1$; $M(\zeta^{(m_1)}) \leqq 1$; $M(\tau^{(m_1)}) \leqq 1$;

$$(17) \qquad W(\eta^{(n_1)}, \zeta^{(m_1)}, \tau^{(m_1)}) > 2.$$

Let $d = d_1$ correspond to $m = m_1$ and $n = n_1$ as stated in Lemma 1. It follows that, if (η, ζ, τ) are eventually defined so that, in addition to satisfying (16), $M(\zeta - \zeta^{(m_1)}) \leqq d_1$ and $M(\tau - \tau^{(m_1)}) \leqq d_1$, then

$$(18) \qquad W(\eta^{(n_1)}, \zeta, \tau) > 1.$$

At this point we have completed the first stage of the definition of (η, ζ, τ). To proceed with the second stage, we define $c_{2j} = d_1 c_{1j}$. Since each component matrix of A is limited, there exist bounds $R_1 > 0$, $F_1 > 0$, and $G_1 > 0$ such that, if (ξ, ρ, σ) are in H_1, then

$$(19) \qquad | W(\xi^{(n_1)}, \rho, \sigma) | \leqq R_1; \quad | W(\xi, \rho^{(m_1)}, \sigma) | \leqq F_1; \quad | W(\xi, \rho, \sigma^{(m_1)}) | \leqq G_1.$$

In view of (12), there exist vectors $(\alpha_2, \beta_2, \gamma_2)$ in H_1 such that

$$W(\alpha_2, \beta_2, \gamma_2) > (3 + 2R_1 + 2F_1 + 2G_1)/c^3_{22}.$$

Let $\xi_2 = c_{22}\alpha_2$; $\rho_2 = c_{22}\beta_2$; $\sigma_2 = c_{22}\gamma_2$; then, there exist integers $m_2 > m_1$ and $n_2 > n_1$ such that

$$W(\xi_2{}^{(n_2)}, \rho_2{}^{(m_2)}, \sigma_2{}^{(m_2)}) > 3 + 2R_1 + 2F_1 + 2G_1. \tag{20}$$

Let us extend the definition of (η, ζ, τ) as follows, where (x_{i2}, r_{j2}, s_{j2}) are, respectively, coördinates of $(\xi_2, \rho_2, \sigma_2)$:

$$y_i = x_{i2}, \ (n_1 < i \leqq n_2); \quad z_j = r_{j2}, \ t_j = s_{j2}, \ (m_1 < j \leqq m_2). \tag{21}$$

We note that $M(\eta^{(n_2)}) \leqq (c^2_{11} + c^2_{22})^{1/2} \leqq (c^2_{11} + c^2_{12})^{1/2} \leqq 1$. Similarly, $M(\zeta^{(m_2)}) \leqq 1$ and $M(\tau^{(m_2)}) \leqq 1$. Let $P_{ijk} = a_{ijk}(y_i z_j t_k - x_{i2} r_{j2} s_{k2})$. Then, we verify that

$$\begin{aligned} W(\eta^{(n_2)}, \zeta^{(m_2)}, \tau^{(m_2)}) &- W(\xi_2{}^{(n_2)}, \rho_2{}^{(m_2)}, \sigma_2{}^{(m_2)}) \\ &= \sum_{i=1}^{n_1} \sum_{j,k=1}^{m_2} P_{ijk} + \sum_{i=n_1+1}^{n_2} \sum_{j=1}^{m_1} \sum_{k=1}^{m_2} P_{ijk} + \sum_{i=n_1+1}^{n_2} \sum_{j=m_1+1}^{m_2} \sum_{k=1}^{m_1} P_{ijk}. \end{aligned} \tag{22}$$

From the definitions of R_1, F_1, and G_1, it is seen that the absolute value of the right member of (22) is at most $(2R_1 + 2F_1 + 2G_1)$. Hence, by use of (20) it follows from (22) that

$$W(\eta^{(n_2)}, \zeta^{(m_2)}, \tau^{(m_2)}) > 3. \tag{23}$$

Let $d = d_2$ correspond to $m = m_2$ and $n = n_2$ thru Lemma 1. It follows that, if (η, ζ, τ) are defined as points in H_1 subject to (16), (21), and to the conditions $M(\zeta - \zeta^{(m_2)}) \leqq d_2$ and $M(\tau - \tau^{(m_2)}) \leqq d_2$, then

$$W(\eta^{(n_2)}, \zeta, \tau) > 2. \tag{24}$$

We have now completed the second stage of the definition of (η, ζ, τ). Each succeeding stage depends on the preceding one as the second depended on the first. For instance, to commence the third stage, we define $c_{3j} = d_2 c_{2j}$; R_2, F_2, and G_2, etc. Then, there exist vectors $(\alpha_3, \beta_3, \gamma_3)$ in H_1 such that

$$W(\alpha_3, \beta_3, \gamma_3) > (4 + 2R_2 + 2F_2 + 2G_2)/c^3_{33};$$

we define $(\xi_3, \rho_3, \sigma_3)$, and integers $m_3 > m_2$ and $n_3 > n_2$. We extend the definition of (η, ζ, τ), and obtain a $d_3 \leqq 1$ from Lemma 1 so that, finally

$$W(\eta^{(n_3)}, \zeta, \tau) > 3, \tag{25}$$

provided that ζ and τ are eventually defined so that $M(\zeta - \zeta^{(m_3)}) \leqq d_3$ and $M(\tau - \tau^{(m_3)}) \leqq d_3$. The continuation of this process thru its successive stages yields a sequence of values (n_i, m_i, d_i) and a well determined set of coördi-

nates for (η, ζ, τ). We remark that these vectors are in H_1 because, at the i-th stage of their definition, we have

$$M(\eta^{(n_i)}) \leqq (c^2_{11} + c^2_{22} + \cdots + c^2_{ii})^{1/2} \leqq (c^2_{11} + c^2_{12} + \cdots + c^2_{1i})^{1/2} \leqq 1,$$

and, similarly, $M(\zeta^{(m_i)}) \leqq 1$ and $M(\tau^{(m_i)}) \leqq 1$. Moreover, for every i,

$$M^2(\zeta - \zeta^{(m_i)}) \leqq \sum_{j=i+1}^{\infty} M^2(\xi_j) \leqq \sum_{j=i+1}^{\infty} c^2_{jj} \leqq d_i^2 \sum_{j=i+1}^{\infty} c^2_{ij} \leqq d_i^2,$$

and, similarly, $M(\tau - \tau^{(m_i)}) \leqq d_i$. Consequently, the successive inequalities (18), (24), (25), etc., are valid:

$$W(\eta^{(n_i)}, \zeta, \tau) > i \qquad (i = 1, 2, \cdots,).$$

Hence, $W(\eta, \zeta, \tau)$ diverges. This contradicts our original hypothesis concerning $W(\xi, \rho, \sigma)$ and completes the proof of Property X.

If we assume that a complete theory for $(m-1)$-linear forms has been developed, we could establish the analogue of Property X for m-linear forms by use of the method just employed for tri-linear forms. In the discussion of the m-linear case we would refer to $(m-1)$-linear forms in the places occupied by bi-linear forms in the proof of Property X in this paper.

UNIVERSITY OF MINNESOTA,
FEBRUARY 15, 1930.

The Constants of the Disturbing Function.

By K. P. Williams.

1. Newcomb's development of the disturbing function requires the calculation of quantities $D^m c_n^{(i)}$, where

$$D = s(d/ds), \qquad c_n^{(i)} = s^{(n-1)/2} b_n^{(i)},$$

and the $b_n^{(i)}$ are the coefficients in the expansion

$$(1 + s^2 - 2s\cos\theta)^{-n/2} = (1/2)b_n^{(0)} + b_n^{(1)}\cos\theta + b_n^{(2)}\cos 2\theta + \cdots$$

where $n = 1, 3, 5, 7$.* In order to obtain the required derivatives Newcomb introduced a general set of quantities $c_n^{(i,j)}$, $j = 1, 2, \cdots, 6$, $c_n^{(i,0)} = c_n^{(i)}$. The quantity $D^m c_n^{(i)}$ is a linear function, with positive integral coefficients, of $c_n^{(i)}, c_n^{(i,1)}, \cdots, c_n^{(i,m)}$. The computation problem therefore reduces to the determination of the $c_n^{(i,j)}$. Newcomb's method of obtaining these numbers has been commended for its accuracy. It is, however, long, although addition and subtraction logarithms were employed to compute the continued fractions that are the foundation of the method. Apparently the same process had to be carried out for each of the values of n.† Newcomb's failure to discover certain remarkable recursion relations that allow the $c_3^{(i,j)}$ to be obtained from the $c_1^{(i,j)}$, the $c_5^{(i,j)}$ from the $c_3^{(i,j)}$, etc., both rapidly and accurately, probably came from his method of handling the quantities. He introduced them as hypergeometric functions with rather complicated parameters. He then apparently looked through the list of relations given by Gauss to find formulas that might be useful. Such a method will evidently fail to reveal relations that owe their origin to the special parameters involved.

2. In this paper a rapid and accurate method of computing the coefficients used by Newcomb will be developed. It is adapted throughout to machine computation. The recursion relations that are used will be obtained by contour integration. This seems the modern way of dealing with the

* For simplicity of typing the manuscript the letter s was used in place of the customary a, that is, $s = a/a'$, where a is the semi-major axis of the orbit of the inner planet and a' that of the outer.

† Newcomb's exposition is given in *Astronomical Papers of the American Ephemeris*, Vol. III, Chapters II and III. Tables of the constants for certain pairs of planets will be found in the last memoir of that volume, and in Vol. V of the same papers, pp. 339-348.

problem. By breaking up the integrand in various ways, integrating by parts, etc., one can hope to discover all important relations connecting the numbers being considered. Some of the relations employed are already in the literature, but they will be developed in order to give completeness to the method. For the same reason the relation that is used to express $D^m c_n^{(i)}$ in terms of the $c_n^{(i,j)}$ will be obtained by integration, although it is not required in the actual problem. Those of the $c_1^{(i,j)}$ that are calculated directly are found by series identical with or similar to series recently given by Brown.

I. THE FUNDAMENTAL EXPRESSIONS.

3. Let

(1) $$G = 1 + s^2 - 2s\cos\theta,$$

and put

$$z = e^{\iota\theta}, \qquad \iota = (-1)^{1/2}$$

so that

$$
\begin{aligned}
(2^{i}) \qquad G &= 1 + s^2 - s(z + 1/z) \\
(2^{ii}) \qquad &= (1/z - s)(z - s) \\
(2^{iii}) \qquad &= (z - s)(1 - sz)/z \\
(2^{iv}) \qquad &= -s(z - 1/s)(z - s)/z.
\end{aligned}
$$

We have from (2^{iv})

$$G^{-n/2} = \frac{1}{(-s)^{n/2}}\left[\frac{z}{(z - 1/s)(z - s)}\right]^{n/2}.$$

It is seen that $G^{-n/2}$ has branch points at $z = 0$, s, and $1/s$, the latter two points being also infinities. It is also seen that the function is single valued in the ring between $|z| = s$ and $|z| = 1/s$, so that within this region it can be expanded in a Laurent series. Therefore

$$G^{-n/2} = (1/2)\sum_{i=-\infty}^{\infty} b_n^{(i)} z^i.$$

Since

$$G^{-n/2}(1/z) = G^{-n/2}(z),$$

it follows that $b_n^{(-i)} = b_n^{(i)}$. We can accordingly write

$$G^{-n/2} = (1/2)\sum_{i=0}^{\infty} b_n^{(i)} z^i + (1/2)\sum_{i=1}^{\infty} b_n^{(i)} z^{-i}.$$

The expansion being valid on the unit circle we have

(3) $$
\begin{aligned}
(1 + s^2 - 2s\cos\theta)^{-n/2} &= (1/2)\sum_{i=0}^{\infty} b_n^{(i)} e^{\iota i\theta} + (1/2)\sum_{i=1}^{\infty} b_n^{(i)} e^{-\iota i\theta} \\
&= (1/2) b_n^{(0)} + \sum_{i=1}^{\infty} b_n^{(i)} \cos i\theta.
\end{aligned}
$$

In accordance with the theorem of Laurent

$$(4) \qquad b_n^{(i)} = (1/\pi\iota)\int z^{-i-1}G^{-n/2}dz,$$

where the integral may be taken around the unit circle. Since $b_n^{(-i)} = b_n^{(i)}$ we have

$$(4') \qquad b_n^{(i)} = (1/\pi\iota)\int z^{i-1}G^{-n/2}dz.$$

4. In order to obtain ultimately the $c_n^{(i,j)}$ of Newcomb we introduce the general set of quantities

$$(5) \qquad g(i,j,n) = \int z^{i-1}(z-s)^jG^{-j-n/2}dz,$$

the path being the circle $|z|=1$. The dependence of g upon s need not be exhibited.

The function g is expressible in terms of the hypergeometric function, and the relation between the two will be obtained for the case $n=1$.

We have, using (2‴),

$$g(i,j,1) = \int z^{j+i-1/2}(z-s)^{-1/2}(1-sz)^{-j-1/2}dz,$$

or, putting $z=st$,

$$g(i,j,1) = s^{j+1}\int t^{j+i-1/2}(t-1)^{-1/2}(1-s^2t)^{-j-1/2}dt,$$

the path being the circle $t=1/s$. The path can be changed into:

(1) The real axis from $t=1/s$ to $t=1+\epsilon$,
(2) A semi-circle of radius ϵ about $t=1$, above the real axis,
(3) The real axis from $t=1-\epsilon$ to $t=\epsilon$,
(4) A circle about $t=0$,
(5) A line from $t=\epsilon$ to $t=1-\epsilon$,
(6) A semi-circle of radius ϵ about $t=1$ and below the real axis.
(7) The real axis from $t=1+\epsilon$ to $t=1/s$.

It is easily shown that the portions of the integral from (2), (4), and (6) approach 0 as $\epsilon\to 0$, that the portions from (1) and (7) cancel each other, and that those from (3) and (5) are equal. Therefore, since the description of (2) leads us to replace $(t-1)^{-1/2}$ by

$$e^{-\pi\iota/2}(1-t)^{-1/2} = -\iota(1-t)^{-1/2}$$

we have

$$g(i,j,1) = 2\iota s^{j+1}\int_0^1 t^{j+i-1/2}(1-t)^{-1/2}(1-s^2t)^{-j-1/2}dt.$$

But it is well known that

$$\int_0^1 u^{b-1}(1-u)^{c-b-1}(1-uz)^{-a}du = \frac{\Gamma(b)\Gamma(c-b)}{\Gamma(c)} F(a, b, c, z)$$

where Γ is the ordinary gamma function and F is the hypergeometric function.*

It follows that

$$g(i, j, 1) = 2\iota s^{j+i} \frac{\Gamma(j+i+1/2)\Gamma(1/2)}{\Gamma(j+i+1)} \times F(j+1/2, j+i+1/2, j+i+1, s^2).$$

Using the relations

$$\Gamma(x+1) = x\Gamma(x), \quad \Gamma^2(1/2) = \pi, \quad \Gamma(j+i+1) = (j+1)\,!,$$

we find

(6) $$g(i, j, 1) = \frac{\pi\iota s^{j+i}[1\cdot 3\cdots(2j+2i-1)]}{2^{j+i-1}(j+i)\,!} \times F(j+1/2,\ j+i+1/2,\ j+i+1,\ s^2).$$

5. Let us now put

(7) $$c_n^{(i,j)} = L(j, n)\, g(i, j, n),$$

where

(8) $$L(0, n) = (1/\pi\iota)\, s^{(n-1)/2}, \qquad L(j, n) = (1/\pi\iota)\, n(n+2)\cdots(n+2j-2)\, s^{(n+2j-1)/2}.$$

For simplicity we set

(9) $$c_n^{(i)} = c_n^{(i,0)}.$$

We have

$$c_n^{(i)} = (1/\pi\iota)\, s^{(n-1)/2} \int z^{i-1} G^{-n/2} dz,$$

so that

(10) $$c_n^{(i)} = s^{(n-1)/2} b_n^{(i)}.$$

In particular,

(11) $$c_1^{(i)} = b_1^{(i)}.$$

II. THE COMPUTATION OF SERIES FOR THE $c_1^{(i,j)}$.

6. The present method requires the direct calculation of certain of the $c_1^{(i,j)}$. Brown has given rapidly, converging series for $c_1^{(10)}$ and $c_1^{(11)}$†; these with a well known recursion relation will give the remaining $c_1^{(i)}$. We need

---

* Whittaker and Watson, *Modern Analysis*, 3rd Ed., p. 293.

† *Monthly Notices of the Royal Astronomical Society*, Vol. 88, pp. 459-469. As a matter of fact we use a series for $c_1^{(9)}$ in place of Brown's series for $c_1^{(11)}$.

also the values of $c_1^{(10,j)}$, $j=1, 2, \cdots, 6$. Series similar to those of Brown will be obtained for $j=1, 2, 3$; recursion relations will be used for $j=4, 5, 6$.

7. From (6) and (8) we have

(12)
$$c_1^{(i,j)}=\frac{[1\cdot 3\cdots(2j-1)]\,[1\cdot 3\cdots(2j+2i-1)]}{2^{j+i-1}(j+i)\,!}\,s^{2j+i}\,F(a,b,c,s^2),$$

where

(13) $$a=j+1/2,\quad b=j+i+1/2,\quad c=j+i+1.$$

It follows from the differential equation for the hypergeometric function that $F(a,b,c,s^2)$ satisfies the equation

(14) $$s(s^2-1)(d^2F/ds^2)+[(2a+2b+1)s^2-1-2c](dF/ds)+4absF=0.$$

Let us now make the change of variables

(15) $$s^2=y/(1+y).$$

We then have from (12)

(16)
$$c_1^{(i,j)}=\frac{2^j[1\cdot 3\cdots(2j-1)]\,[1\cdot 3\cdots(2j+2i-1)]}{(j+i)\,!}\left(\frac{s}{2}\right)^{2j+i-1}y^{1/2}H,$$

where

(17) $$H(a,b,c,y)=(1+y)^{-1/2}\,F\{a,b,c,y/(1+y)\}.$$

Using (14) it is found that H satisfies the equation

$$4y(1+y)^3(d^2H/dy^2)+4(1+y)^2\,[c+(2+c-a-b)y]\,(dH/dy)+(1+y)\,[2c-4ab+(1+2c-2a-2b)y]\,H=0.$$

If we make use of (13) and denote H when expressed in terms of i, j, and y by $H(i,j,y)$ we find

(18) $$4y(1+y)^2H''(i,j,y)+4(1+y)\,[i+j+1+(2-j)y]H'(i,j,y)+[1-2j-4ij-4j^2+(1-2j)y]H(i,j,y)=0.$$

8. The equation (18) has the one singular point $y=-1$, but is regular at that point. It appears from (17) that the value of H desired is the expansion about $y=0$ in which the constant term has the value 1. The equation (18) is satisfied by the series

(19) $$1+a_1y+a_2y^2+\cdots$$

where, in particular,

$$a_1=(2j+4ij+4j^2-1)/4(i+j+1).$$

The value of y in terms of s is

(20) $$y = s^2/(1 - s^2).$$

Thus y is positive since $0 < s < 1$. But $y > 1$ if $s > 2^{1/2}/2$. Since the series (19) does not converge for $y > 1$, on account of the singular point $y = -1$, it is necessary, in certain cases, to use the analytic continuation of the series. This is easily done; the following considerations will govern the development.

The greatest value of s for the major planets is for Venus-Earth, for which $s = .723$, giving $y = 1.097$. We will derive an expansion for H about $y = .6$. Such an expansion will converge for $|y - .6| < 1.6$, since the only singular point of (16) is $y = -1$.

Putting $t = y - .6$, we write (18) in the form

(21) $$\begin{aligned} &4(.6 + t)(1.6 + t)^2 H''(i, j, t) \\ &+ 4(1.6 + t)[i + j + 1 + .6(2 - j) + (2 - j)t]H'(i, j, t) \\ &+ [1 - 2j - 4ij - 4j^2 + .6(1 - 2j) + (1 - 2j)t]H(i, j, t) = 0. \end{aligned}$$

This equation is satisfied by

$$H = b_0 + b_1 t + b_2 t^2 + \cdots,$$

where b_0 is obtained by substituting $y = .6$ in (19) and b_1 by substituting $y = .6$ in

$$a_1 + 2a_2 y + 3a_3 y^2 + \cdots.$$

The other coefficients are determined in terms of b_0 and b_1.

The actual series for the values of i and j required are given in V. It will be seen that they converge rapidly.

III. THE RECURSION RELATIONS.

9. In order to obtain the recursion relations for the $c_n^{(i,j)}$ we establish such relations for the functions $g(i, j, n)$.

We begin by deducing by contour integration a well known relation connecting the values of $c_1^{(i)}$ for three successive values of i.

We have from (5)

$$g(i, 0, 1) = \int z^{i-1} G^{-n/2} dz.$$

Observing from (2^i) that

$$\partial G/\partial z = - s + s/z^2,$$

we see that we can write

$$g(i, 0, 1) = -(1/s) \int z^{i-1} G^{-n/2} (\partial G/\partial z)\, dz + g(i - 2, 0, 1).$$

If we integrate by parts, observing that on account of the single valued property of $G^{-n/2}$ along the path of integration, the integrated term in the familiar formula is zero, we obtain readily

$$g(i,0,1)=\frac{2i-2}{2i-1}(s+1/s)g(i-1,0,1)-\frac{2i-3}{2i-1}g(i-2,0,1).$$

From (7) and (8) it follows that

(22)
$$c_1^{(i)}=\frac{2i-2}{2i-1}(s+1/s)c_1^{(i-1)}-\frac{2i-3}{2i-1}c_1^{(i-2)}.$$

10. We have

$$g(i,j+2,n-2)=\int z^{i-1}(z-s)^{j+2}G^{-j-1-n/2}dz$$
$$=\int z^{i}(z-s)^{j+1}G^{-j-1-n/2}dz-s\int z^{i-1}(z-s)^{j+1}G^{-j-1-n/2}dz,$$

from which we obtain

(23) $$g(i+1,j+1,n)=sg(i,j+1,n)+g(i,j+2,n-2).$$

If we multiply the two members of this relation by $L(j+2,n-2)$, and observe that

(24) $$L(j+2,n-2)=(n-2)L(j+1,n),$$

we find

$$c_n^{(i+1,j+1)}=sc_n^{(i,j+1)}+\frac{1}{n-2}c_{n-2}^{(i,j+2)},$$

or, changing i into $i-1$, and j into $j-1$,

(25) $$c_n^{(i,j)}=sc_n^{(i-1,j)}+\frac{1}{n-2}c_{n-2}^{(i-1,j+1)}.$$

11. From (2^{ii}) it follows that

(26) $$G^{-j-n/2}=G^{-j-1-n/2}(1/z-s)(z-s),$$

from which it is seen that

(27) $$g(i,j,n)=g(i-1,j+1,n)-sg(i,j+1,n).$$

Multiplying by $L(j+1,n)$ and observing that

$$L(j+1,n)=(n+2j)sL(j,n),$$

we obtain

(28) $$c_n^{(i-1,j+1)}=s[c_n^{(i,j+1)}+(n+2j)c_n^{(i,j)}].$$

12. Replacing i by $i-1$ in (27) we have

$$(29)\qquad g(i-1, j, n) = g(i-2, j+1, n) - sg(i-1, j+1, n).$$

On the other hand we have by integrating by parts, putting

$$dv = (z-s)^j dz,$$

the result

$$g(i, j, n) = -\frac{i-1}{j+1}\int (z-s)^{j+1} z^{i-2} G^{-j-n/2} dz$$
$$- \frac{s(n+2j)}{2(j+1)}\int (z-s)^{j+1} z^{i-1}(1-1/z^2) G^{-j-1-n/2} dz$$

$$= -\frac{i-1}{j+1}\int z^{i-1}(z-s)^j G^{-j-n/2} dz + \frac{i-1}{j+1} s \int z^{i-2}(z-s)^j G^{-j-n/2} dz$$
$$- \frac{s(n+2j)}{2(j+1)}\int z^{i-1}(z-s)^{j+1} G^{-j-1-n/2} dz$$
$$+ \frac{s(n+2j)}{2(j+1)}\int z^{i-3}(z-s)^{j+1} G^{-j-1-n/2} dz$$
$$= -\frac{i-1}{j+1} g(i, j, n) + \frac{(i-1)s}{j+1}\, g(i-1, j, n)$$
$$- \frac{s(n+2j)}{2(j+1)}\, g(i, j+1, n) + \frac{s(n+2j)}{2(j+1)}\, g(i-2, j+1, n).$$

From this we derive

$$(30)\qquad 2(j+1) g(i, j, n) - 2(i-1) sg(i-1, j, n)$$
$$= (n+2j) sg(i-2, j+1, n) - (n+2j) sg(i, j+1, n).$$

Eliminating $g(i-2, j+1, n)$ between (29) and (30) and then eliminating $g(i-1, j+1, n)$ by means of (27), we have

$$(31)\qquad (n+2j+2i-2) sg(i-1, j, n) - [2(j+i) - (n+2j)s^2] g(i, j, n)$$
$$= (n+2j) s(1-s^2) g(i, j+1, n).$$

Multiplying (31) by $L(j+1, n)$ and afterwards replacing j by $j+1$, we have

$$(32)\qquad (n+2j+2i) sc_n^{(i-1, j+1)} - [2(j+i+1) - (n+2j+2)s^2] c_n^{(i, j+1)}$$
$$= (1-s^2) c_n^{(i, j+2)}.$$

12. In order to obtain the next relation it is convenient to set

$$A = g(i+1, j, n), \qquad X = g(i-1, j+1, n),$$
$$B = g(i, j+1, n-2), \qquad Y = g(i, j+1, n),$$
$$C = g(i, j+2, n-2), \qquad Z = g(i+1, n+1, n).$$

Using (26) it is seen that

(33) $$A = Y - sZ.$$

On the other hand

(34) $$C = -sY + Z.$$

If we integrate by parts we obtain

$$B = -\frac{j+1}{i} A - \frac{s(n+2j)}{2i} Z + \frac{s(n+2j)}{2i} X,$$

so that

(35) $$2iB + 2(j+1)A = s(n+2j)(X - Z).$$

But if we put

$$G^{-j-n/2} = (1 + s^2 - sz - s/z) G^{-j-1-n/2},$$

we find

(36) $$B = (1+s^2)Y - sZ - sX.$$

Eliminating X between (35) and (36) we obtain

$$(n+2j+2i)B + 2(j+1)A = (n+2j)[(1+s^2)Y - 2sZ].$$

But from (33) and (34)

$$A - sC = (1+s^2)Y - 2sZ,$$

so that

$$(n+2j+2i)B + 2(j+1)A = (n+2j)(A - sC).$$

We thus obtain finally

(37) $$(n-2)g(i+1, j, n) = (n+2j+2i)g(i, j+1, n-2) + s(n+2j)g(i, j+2, n-2).$$

If we multiply by $L(j+2, n-2)$ and use (24) as well as the relation

$$L(j+2, n-2) = (n+2j)sL(j+1, n-2)$$

we find

(38) $$(n-2)^2 c_n^{(i+1, j)} = (n+2j+2i)c_{n-2}^{(i, j+1)} + c_{n-2}^{(i, j+2)},$$

a relation obtained by Innes in a different way.*

If we eliminate $c_n^{(i-1, j+1)}$ between (28) and (32) we find

(39) $$c_n^{(i, j+2)} = (n+2j)(n+2j+2i)[s^2/(1-s^2)]c_n^{(i, j)} - 2\{(n+2j+i+1) - (n+j)[1/(1-s^2)]\}c_n^{(i, j+1)}.$$

---

* *Monthly Notices of the Royal Astronomical Society*, Vol. 69, p. 639. The relation is however printed erroneously.

All the necessary recursion relations have now been obtained.

13. Although the aim of the present paper is to give a convenient method for computing the $c_n^{(i,j)}$, we shall for the sake of completeness, as stated in § 2, obtain the basic formula by which the $D^m c_n^{(i)}$ are expressed in terms of the $c_n^{(i,j)}$.

We have

$$s(\partial g/\partial s) = -js\int z^{i-1}(z-s)^{j-1}G^{-j-n/2}dz$$
$$-(j+n/2)s\int z^{i-1}(z-s)^{j}G^{-j-1-n/2}(\partial G/\partial s)\,dz.$$

But

$$\partial G/\partial s = -2(z-s)-(z/s)\partial G/\partial z.$$

If we use this expression and integrate by parts we find

(40) $$s(\partial g/\partial s) = (i+j)g(i,j,n)+(n+2j)sg(i,j+1,n).$$

Put

$$D = s(\partial/\partial s);$$

then from (7) we have

(41) $$Dc_n^{(i,j)} = g(i,j,n)DL(j,n)+L(j,n)Dg(i,j,n).$$

If now we multiply (40) by $L(j,n)$ and use the relation

$$(n+2j)sL(j,n) = L(j+1,n)$$

to transform it we find by (41), since

$$DL(j,n) = [(n+2j-1)/2]\,L(j,n),$$

the ultimate relation

(42) $$Dc_n^{(i,j)} = [(n+2i+4j-1)/2]\,c_n^{(i,j)} + c_n^{(i,j+1)}.$$

We have for instance

$$Dc_n^{(i)} = [(n+2i-1)/2]\,c_n^{(i)} + c_n^{(i,1)}$$
$$D^2c_n^{(i)} = [(n+2i-1)/2]^2c_n^{(i)} + (n+2i+1)c_n^{(i,1)} + c_n^{(i+2)},$$

etc.*

* It is not desirable to give general algebraic expressions for $D^m c_n^{(i)}$ for larger values of m, but merely to compute the numerical coefficients. A table of the coefficients is given by Newcomb, *loc. cit.*, V., p. 313.

IV. OUTLINE OF THE COMPUTATIONS.

14. To compute the $c_1^{(i,j)}$.*

The plan is to compute each column of the array starting from the bottom.

(1) From the value of s compute the values of

$$(43) \qquad \begin{aligned} y &= s^2/(1-s^2), \qquad & s_1 &= 1/(1-s^2), \\ s' &= s + 1/s, \qquad & s_2 &= s/(1-s^2). \end{aligned}$$

(2) From the series expansions compute the values of $c_1^{(9)}$ and $c_1^{(10)}$.

(3) From (22) we have

$$(44) \qquad c_1^{(i-2)} = \frac{2i-2}{2i-3}\, s' c_1^{(i-1)} - \frac{2i-1}{2i-3} c_1^{(i)}.$$

Beginning with the values of $c_1^{(10)}$ and $c_1^{(9)}$ compute successively $c_1^{(8)}, \cdots$, and $c_1^{(0)}$.†

(4) Using the series expansions compute $c_1^{(10,1)}$, $c_1^{(10,2)}$, $c_1^{(10,3)}$.

(5) From (28) we have, for $n=1$,

$$(45) \qquad c_1^{(i-1,j+1)} = s(c_1^{(i,j+1)} + (1+2j)c_1^{(i,j)}).$$

This formula enables us to complete the second, third, and fourth columns, by working from the bottom upwards.‡ The computation is made to include the case $i=0$, so as to give $c_1^{(-1,1)}$, $c_1^{(-1,2)}$, $c_1^{(-1,3)}$. Such quantities were not employed by Newcomb, but are needed here in order to start the table of the $c_3^{(i,j)}$.

(6) The remaining columns are found in a similar way except that the bottom elements are gotten by recursion relations instead of from series. From (32) we find, on putting $n=1$ and $j=2, 3, 4$, successively,

$$(46) \qquad \begin{aligned} c_1^{(10,4)} &= 25 s_2 c_1^{(9,3)} - (26 s_1 - 7y) c_1^{(10,3)}, \\ c_1^{(10,5)} &= 27 s_2 c_1^{(9,4)} - (28 s_1 - 9y) c_1^{(10,4)}, \\ c_1^{(10,6)} &= 29 s_2 c_1^{(9,5)} - (30 s_1 - 11y) c_1^{(10,5)}. \end{aligned}$$

(7) We also need $c_1^{(-1,7)}$. From (39) we have, putting $n=1$, $i=-1$, $j=5$,

* It is assumed that the $c_1^{(i,j)}$ are needed up to $i=10$, $j=6$. More rows and columns can be added if necessary. For small values of s the entire computation can be shortened by methods that will suggest themselves.

† Since $c_1^{(0)}$ is an elliptic integral of the first kind, it can be computed by a rapidly converging theta series, thus checking the first column.

‡ If we solve (45) for $c_1^{(i,j+1)}$ we have a means, in conjunction with (44), to work downwards, adding new rows.

(47) $$c_1^{(-1,7)} = 99yc_1^{(-1,5)} - (22 - 12s_1)c_1^{(-1,6)}.$$

15. To calculate the $c_3^{(i,j)}$.

(1) Putting $n = 3$, $i = -1$ in (38) we have

(48) $$c_3^{(0,j)} = (1 + 2j)c_1^{(-1,j+1)} + c_1^{(-1,j+2)},$$

which gives the first row in the table up to $j = 5$.

(2) Putting $n = 3$ in (25) we have

(49) $$c_3^{(i,j)} = sc_3^{(i-1,j)} + c_1^{(i-1,j+1)},$$

which allows the table to be extended to as great a value of i as desired.

(3) We need the values of $c_3^{(-1,j)}$ in order to start the table of the $c_5^{(i,j)}$. From (28) we have

(50) $$c_3^{(-1,j+1)} = s[c_3^{(0,j+1)} + (3 + 2j)c_3^{(0,j)}].$$

This will give the values from $c_3^{(-1,1)}$ to $c_3^{(-1,5)}$. We also require $c_3^{(-1,6)}$. From (39)

(51) $$c_3^{(-1,6)} = 99yc_3^{(-1,4)} - (22 - 14s_1)c_3^{(-1,5)}.$$

16. The calculation of the $c_5^{(i,j)}$ is similar to that of the $c_3^{(i,j)}$, etc.

It is seen that there will be one fewer columns in the array of the $c_3^{(i,j)}$ than in that of the $c_1^{(i,j)}$. If necessary, a new column can be added to the $c_3^{(i,j)}$ as the last columns were added in the case of the $c_1^{(i,j)}$, calculating the final element of the column by (32) and working up the column by (28).

The recursion formulas that form the distinctive feature of the present method are (25) and (28). They are the only ones that are repeatedly used with the exception of the classic formula (22). These formulas possess in a marked degree the characteristics essential for computational purposes. They do not lead to an accumulation of errors for coefficients are small and there are no negative terms, and furthermore they are rapidly used. Formula (25) is particularly accurate for $n = 5, 7$. The coefficients in (25) do not depend upon either i or j, while those in (28) are constant down a column. This is a great advantage. Formula (32) does not have the merit of (25) and (28). This is seen by inspecting (46). Errors accumulate through its repeated use. On this account it is desirable to compute $c_1^{(10,3)}$ to two more places that will be retained in $c_1^{(10,4)}$ etc. The rapidity of convergence of the series makes this practical. A similar criticism can be made of (39). No extensive use of it is made, however.* For smaller values of i, it may happen

* Newcomb's method of obtaining approximate values for numbers in columns to the right of those accurately computed was solely one of induction.

that (39) gives $c_n^{(i,j)}$ as the sum and not the difference, of two quantities, and the formula becomes more accurate.*

V. THE NUMERICAL SERIES.

17. The series employed in § 14 (2) and (4), and whose derivation was explained in II follow.† The series in y are used if $y \leqq .3$, otherwise the series in $t = y - .6$.

$c_1^{(9)} = 4 \cdot 5 \cdot 11 \cdot 13 \cdot 17 \cdot 19 (s/4)^8 y^{1/2}$ multiplied by one of the series

$1.000\ 000\ 00$	$0.985\ 836\ 16$
$-0.025\ 000\ 00y$	$-0.022\ 336\ 71t$
$+0.002\ 556\ 82y^2$	$+0.001\ 933\ 39t^2$
$-0.000\ 443\ 89y^3$	$-0.000\ 270\ 37t^3$
$+0.000\ 104\ 57y^4$	$+0.000\ 049\ 43t^4$
$-0.000\ 030\ 25y^5$	$-0.000\ 010\ 81t^5$
$+0.000\ 010\ 17y^6$	$+0.000\ 002\ 75t^6$
$-0.000\ 003\ 84y^7$	
$+0.000\ 001\ 59y^8$	
$-0.000\ 000\ 71y^9$	

$c_1^{(10)} = 2 \cdot 11 \cdot 13 \cdot 17 \cdot 19 (s/4)^9 y^{1/2}$ multiplied by one of the series

$1.000\ 000\ 00$	$0.987\ 065\ 29$
$-0.022\ 727\ 27y$	$-0.020\ 485\ 38t$
$+0.002\ 130\ 68y^2$	$+0.001\ 643\ 37t^2$
$-0.000\ 341\ 45y^3$	$-0.000\ 214\ 74t^3$
$+0.000\ 074\ 69y^4$	$+0.000\ 036\ 71t^4$
$-0.000\ 020\ 17y^5$	$-0.000\ 006\ 78t^5$
$+0.000\ 006\ 35y^6$	
$-0.000\ 002\ 26y^7$	
$+0.000\ 000\ 88y^8$	

$c_1^{(10,1)} = 16 \cdot 13 \cdot 17 \cdot 19 \cdot 21 (s/4)^{11} y^{1/2}$ multiplied by one of the series

$1.000\ 000\ 00$	$1.544\ 629\ 77$
$+0.937\ 500\ 00y$	$+0.880\ 010\ 95t$
$-0.053\ 485\ 58y^2$	$-0.043\ 047\ 37t^2$
$+0.007\ 136\ 42y^3$	$+0.004\ 720\ 19t^3$
$-0.001\ 385\ 03y^4$	$-0.000\ 726\ 91t^4$
$+0.000\ 340\ 46y^5$	$+0.000\ 137\ 82t^5$
$-0.000\ 099\ 06y^6$	$-0.000\ 030\ 24t^6$
$+0.000\ 032\ 77y^7$	$+0.000\ 007\ 50t^7$
$-0.000\ 012\ 01y^8$	$-0.000\ 002\ 34t^8$
$+0.000\ 004\ 78y^9$	
$-0.000\ 002\ 04y^{10}$	
$+0.000\ 000\ 93y^{11}$	

* For instance, in the case of Venus-Earth, if we determine $c_1^{(1,6)}$ from (39) using Newcomb's values for $c_1^{(1,4)}$ and $c_1^{(1,5)}$ we find precisely his value to all seven places.

† The series in y for $c_1^{(10)}$ is the same as one of the series in Brown's paper except as to the number of terms and decimal places. Brown also gives expansions about $y = .5$ and $y = 1$. All the series here given were computed by Miss Irene Price.

$c_1^{(10,2)} = 32 \cdot 13 \cdot 17 \cdot 19 \cdot 21 \cdot 23 (s/4)^{13} y^{\frac{1}{2}}$ multiplied by one of the series

$1.000\ 000\ 00$	$2.426\ 804\ 77$
$+ 1.903\ 846\ 15y$	$+ 2.832\ 665\ 16t$
$+ 0.825\ 721\ 15y^2$	$+ 0.727\ 879\ 22t^2$
$- 0.064\ 603\ 37y^3$	$- 0.045\ 893\ 56t^3$
$+ 0.010\ 370\ 11y^4$	$+ 0.005\ 842\ 58t^4$
$- 0.002\ 262\ 36y^5$	$- 0.000\ 987\ 77t^5$
$+ 0.000\ 600\ 84y^6$	$+ 0.000\ 199\ 02t^6$
$- 0.000\ 184\ 17y^7$	$- 0.000\ 045\ 05t^7$
$+ 0.000\ 063\ 11y^8$	$+ 0.000\ 009\ 92t^8$
$- 0.000\ 023\ 66y^9$	
$+ 0.000\ 009\ 56y^{10}$	
$- 0.000\ 004\ 12y^{11}$	
$+ 0.000\ 001\ 87y^{12}$	
$- 0.000\ 000\ 89y^{13}$	

$c_1^{(10,3)} = 256 \cdot 17 \cdot 19 \cdot 21 \cdot 23 \cdot 25 (s/4)^{15} y^{\frac{1}{2}}$ multiplied by one of the series *

$1.000\ 000\ 00$	$3.824\ 315\ 72$
$+ 2.875\ 000\ 00y$	$+ 6.768\ 654\ 73t$
$+ 2.653\ 125\ 00y^2$	$+ 3.801\ 102\ 49t^2$
$+ 0.701\ 318\ 36y^3$	$+ 0.583\ 328\ 93t^3$
$- 0.062\ 099\ 50y^4$	$- 0.039\ 025\ 67t^4$
$+ 0.011\ 030\ 56y^5$	$+ 0.005\ 305\ 42t^5$
$- 0.002\ 588\ 22y^6$	$- 0.000\ 941\ 87t^6$
$+ 0.000\ 723\ 77y^7$	$+ 0.000\ 193\ 22t^7$
$- 0.000\ 230\ 04y^8$	$- 0.000\ 033\ 27t^8$
$+ 0.000\ 080\ 82y^9$	
$- 0.000\ 030\ 81y^{10}$	
$+ 0.000\ 012\ 58y^{11}$	
$- 0.000\ 005\ 44y^{12}$	
$+ 0.000\ 002\ 48y^{13}$	
$- 0.000\ 001\ 18y^{14}$	
$+ 0.000\ 000\ 58y^{15}$	

INDIANA UNIVERSITY.

* In spite of the number of integers that occur in the coefficients of the various series the power of $s/4$ reduces it in every case to a small quantity. For example in the case of Venus-Earth, for which $\log s = 9.85\ 9337$, $\log y = 0.04\ 0346$, the coefficient of the series is found to be 0.00 03601.

A System of Six Rectangular Biquadratics.

By Mildred Waters Dean.

This paper discusses a special set of six points in the inversive plane, and certain biquadratic curves related to them. It takes up the following topics:

I. Representing the three vertices of a triangle by complex quantities a, b, and c, and studying the locus traced out by a point moving so as to subtend equal angles (modulo π) at two sides ab and ac of the triangle, we know that this locus is a biquadratic having a double point at a, cutting itself at right angles there, and passing through the other two vertices of the triangle, its two Fermat points (denoted by f and f_1) and the point ∞. Then, in addition to the biquadratic cutting itself at right angles at a, and passing through b, c, f, f_1 and ∞, there are five other curves which bear the same relation to the set of six points a, b, c, f, f_1 and ∞, namely: five biquadratics cutting themselves at right angles at b, c, f, f_1 and ∞, respectively, and passing through the other five points of the set. Thus, the three vertices of the triangle, the Fermat points, and the point ∞ are a special set of six points such that there can be found six biquadratic curves, each having a double point at one of the six points, cutting itself at right angles there, and passing through the other five.

Moreover, it is shown that this set of six points can be broken up into two triads which are apolar.

II. The study of the conditions under which, given a set of six points, the biquadratic having a double point at each and passing through the others will be rectangular. (A rectangular curve is one cutting itself at right angles at a double point.) A set of six points in the inversive plane is chosen, with the restriction that one of them be the point ∞, and the necessary conditions are derived.

III. We next give the geometrical interpretation of the conditions derived in Section II. It is known that five points have a biquadratic invariant I_2.* For any given four points this is a covariant biquadratic curve C_2. The relation of the four and any point on C_2 is symmetrical. It is found that the set of six points mentioned in Section II is such that I_2 vanishes for any five, or, in other words, that the covariant C_2 of any four is on the other two.

---

* This invariant I_2 is given in a "Note on Neuberg's Cubic Curve," by Professor Morley, *American Mathematical Monthly*, Vol. 32 (1925), pp. 407-411.

IV. It is then pointed out that the six points, together with the associated biquadratics, may be taken as the fundamental points and curves of a Geiser involution in the inversive plane.

I.

We have before us the problem of finding the locus of a point which moves so as to subtend equal angles, modulo π, at two sides of a given triangle abc.* (By "equal modulo π" we mean that either the angles are equal, or they differ by π.) If we denote our moving point by a complex quantity z, we may first try to find its locus when it moves so as to subtend equal angles (modulo π) at the sides ab and ac. The line from b to z is a vector $z-b=\rho_1 e^{i\theta_1}$. Likewise, the line from a to z is a vector $z-a=\rho_2 e^{i\theta_2}$ and that from c to z is a vector $z-c=\rho_3 e^{i\theta_3}$. Now we wish the angles azb and cza to be equal, or else to differ by π. If we take the quotient $(z-b)/(z-a)$, we find that it is equal to $(\rho_1/\rho_2)e^{i(\theta_1-\theta_2)}$, where $\theta_1-\theta_2$ is the angle azb. Similarly, $(z-a)/(z-c)$ is equal to $(\rho_2/\rho_3)e^{i(\theta_2-\theta_3)}$, where $\theta_2-\theta_3$ is the angle cza. Now form the quotient

$$[(z-b)/(z-a)]/[(z-a)/(z-c)]=(\rho_1\rho_3/\rho_2{}^2)e^{i(\theta_1+\theta_3-2\theta_2)}).$$

If the angles azb and cza are to be equal, we must have $\theta_1-\theta_2=\theta_2-\theta_3$, or $\theta_1+\theta_3-2\theta_2=0$. If $azb-cza=\pi$, we have $\theta_1-2\theta_2+\theta_3=\pi$. Thus, the exponent of "e" in the quotient $[(z-b)/(z-a)]/[(z-a)/(z-c)]$ either vanishes or reduces to $i\pi$, so that, in either case, the quotient reduces to a real quantity. Therefore, we have as a first fundamental equation:

$$(1) \qquad [(z-b)/(z-a)]/[(z-a)/(z-c)]= \text{ a real quantity.}$$

But, since a real quantity always equals its conjugate, we may write

$$(2) \qquad [(z-b)/(z-a)]/[(z-a)/(z-c)] = [(\bar z-\bar b)/(\bar z-\bar a)]/[(\bar z-\bar a)/(\bar z-\bar c)],$$

$$(3) \qquad (z-a)^2(\bar z-\bar b)(\bar z-\bar c)-(\bar z-\bar a)^2(z-b)(z-c)=0.$$

* This is, of course, an old problem. See, for example, Salmon's "Higher Plane Curves." The problem is also discussed in an article by Professor M. T. Naraniengar, in the *Proceedings of the Edinburgh Mathematical Society*, Vol. 28, p. 73. The curve was called by the Belgian geometers the "focale à noeud." It is identical with the strophoid, which was first considered by Barrow in *Lectiones Geometricae* (1669), p. 69. Further references are: Van Rees, *Correspondance Mathématique de Quetelet*, tome V, p. 361; Lebon, *Journal de Mathématiques Speciales* (1895); Teixera, *Traité des Courbes Speciales Rémarkables*, tome I, p. 30. A history of the problem is given in Loria's *Speciale Algebraische und Transcendente Ebene Kurven* (1902), pp. 66-67, where various additional references are given.

The curve given by (3) is, inversively speaking, a biquadratic passing through ∞; projectively speaking, it is a circular cubic.* We shall think of it as a biquadratic. It is obvious that the curve passes through a, b, and c. Since the two Fermat points subtend equal angles, modulo π, at the sides of the given triangle, it is evident that our curve also passes through these points.

If we consider the point z in a position very near to a, we know, of course, by hypothesis, that angle $azb =$ angle cza. The angles abz and acz have become infinitesimal, and may be considered as equal. Then the angle $baz =$ angle zac; that is, the angles at the vertex a become more nearly equal as z approaches a. This shows that the curve (3), passing through a, bisects the angle at a. In fact, it can be shown that this curve not only bisects the angle at a, but it has a double point there, and cuts itself at right angles at this point. Let us denote this curve by A.

But A is not the only curve which furnishes us with a solution of our problem. For we may also have

$$(4) \qquad (z-c)^2(\bar{z}-\bar{a})(\bar{z}-\bar{b})-(\bar{z}-\bar{c})^2(z-a)(z-b)=0.$$

This is the locus of a point z which moves so as to subtend equal angles. mod. π, at the sides ac and bc. It is a biquadratic on ∞ which passes through c, cuts itself at right angles there, and bisects the angle at c. It also passes through a and b and the two Fermat points. We denote this curve by C. Similarly,

$$(5) \qquad (z-b)^2(\bar{z}-\bar{a})(\bar{z}-\bar{c})-(\bar{z}-\bar{b})^2(z-a)(z-c)=0$$

will be the locus of a point z moving so as to subtend equal angles, mod. π, at the sides bc and ba. It is also a biquadratic passing through a, b, c, f, f_1 and ∞. It cuts itself at right angles at b, and bisects the angle at b. We shall speak of this curve as the curve B.

So we have found three different curves which are solutions of our problem. We thus have a net of biquadratics on the six points a, b, c, f, f_1 and ∞. Now, since our biquadratics have double points at a, b, and c, respectively, and cut themselves at right angles at these points, is it not natural to suppose that we can find three more curves of the net, which have double points at f, f_1 and ∞, respectively, and which cut themselves at right angles at these points? This assumption is found to be true.

No simple method of obtaining the equations of the two biquadratics having double points at f and f_1 has been found, but we may find the equation of the one having a double point at ∞ as follows: the curve will have an

* See explanatory note at the end of this paper.

equation of the form $LA + MB + NC = 0$, where L, M, and N are real constants, which are to be determined. They must be determined in such a way that the curve will cut itself at right angles at ∞. This means that the coefficients of the terms in $z^2\bar{z}^2$, $z^2\bar{z}$, $z\bar{z}^2$ and $z\bar{z}$ must all be zero. The coefficient of $z^2\bar{z}^2$ is zero already. Equating the coefficient of $z^2\bar{z}$ to zero we have

$$(6) \qquad L(2\bar{a} - \bar{b} - \bar{c}) + M(2\bar{b} - \bar{a} - \bar{c}) + N(2\bar{c} - \bar{a} - \bar{b}) = 0.$$

Equating the coefficient of $z\bar{z}^2$ to zero we have

$$(7) \qquad L(b + c - 2a) + M(a + c - 2b) + N(a + b - 2c) = 0.$$

Equating the coefficient of $z\bar{z}$ to zero,

$$(8) \qquad L[2a(\bar{b} + \bar{c}) - 2\bar{a}(b + c)] + M[2b(\bar{a} + \bar{c}) - 2\bar{b}(a + c)] + N[2c(\bar{a} + \bar{b}) - 2\bar{c}(a + b)] = 0.$$

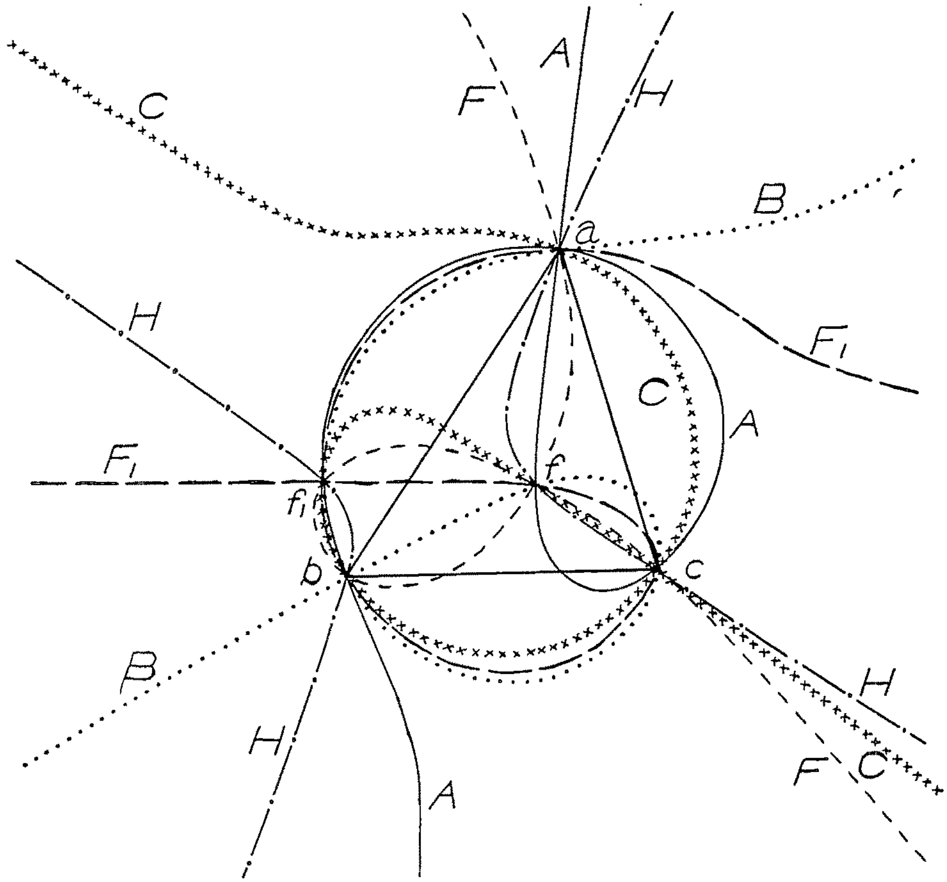
In (6), (7), and (8) we have three linear homogeneous equations in three unknowns L, M, and N. They have solutions different from zero only if

$$\begin{vmatrix} 2\bar{a}-\bar{b}-\bar{c} & 2\bar{b}-\bar{a}-\bar{c} & 2\bar{c}-\bar{a}-\bar{b} \\ b+c-2a & a+c-2b & a+b-2c \\ 2a(\bar{b}+\bar{c})-2\bar{a}(b+c) & 2b(\bar{a}+\bar{c})-2\bar{b}(a+c) & 2c(\bar{a}+\bar{b})-2\bar{c}(a+b) \end{vmatrix} = 0$$

Adding the sum of the second and third columns to the first makes each element of the first column zero. Obviously, then, the values $L = M = N = 1$ will satisfy equations (6), (7), and (8), so that our biquadratic having a double point at infinity and cutting itself at right angles there is $A + B + C = 0$. A biquadratic cutting itself at right angles at infinity is known to be a rectangular hyperbola. In our case, the curve also passes through the three vertices of the triangle and its two Fermat points. The rectangular hyperbola passing through the three vertices of a triangle and its two Fermat points is known as the Kiepert hyperbola of the triangle. So our curve is nothing more than the Kiepert hyperbola of the triangle.

We may now state the following theorem:

THEOREM 1. *The three vertices of a triangle, its two Fermat points, and the point infinity form a set of six points such that the biquadratic having a double point at any one of them, and passing through the others, is rectangular.*



THE SIX RECTANGULAR BIQUADRATICS ASSOCIATED WITH THE TRIANGLE *abc*.

CURVE A—Biquadratic with Double Point at a	————————
CURVE B—Biquadratic with Double Point at b	. . . .
CURVE C—Biquadratic with Double Point at c	x x x x x x x x x x x
CURVE F—Biquadratic with Double Point at f	— — — — — — —
CURVE F_1—Biquadratic with Double Point at f_1	—— —— —— ——
CURVE H—Kiepert Hyperbola	——— . ——— . ———

In addition, these points fall into two triads abc and $ff_1\infty$ which are apolar. If abc and $ff_1\infty$ are to be apolar, we must have

$$3abc-(f+f_1+\infty)(ab+ac+bc) \\ +(a+b+c)(ff_1+f\infty+f_1\infty)-3ff_1\infty=0.$$

Then the sum of all quantities multiplying ∞ must equal zero, or

$$-(ab+ac+bc)+(a+b+c)(f+f_1)-3ff_1 \text{ should equal zero.}$$

Replacing $a+b+c$ by s_1 and $ab+ac+bc$ by s_2 we write:

(9) $$-s_2+s_1(f+f_1)-3ff_1 \text{ should equal zero.}$$

f and f_1 may be expressed in terms of a, b, and c as follows:

$$f=s_1/3-\bar{v}_1v_2/3\bar{v}_2, \qquad f_1=s_1/3-v_1\bar{v}_2/3\bar{v}_1,$$

where $v_1=a+\omega b+\omega^2c$ and $v_2=a+\omega^2b+\omega c$ (ω being one of the complex cube roots of unity.)

We have now

$$f+f_1=2s_1/3-(\bar{v}_1^2v_2+\bar{v}_2^2v_1)/3\bar{v}_1\bar{v}_2,$$
$$ff_1=s_1^2/9+v_1v_2/9-s_1/9(\bar{v}_1^2v_2+\bar{v}_2^2v_1/\bar{v}_1\bar{v}_2).$$

Substituting these values in equation (9) we have

$$-s_2+2s_1^2/3-(s_1/3)[(\bar{v}_1^2v_2+\bar{v}_2^2v_1)/\bar{v}_1\bar{v}_2] \\ -s_1^2/3-v_1v_2/3+(s_1/3)[(\bar{v}_1^2v_2+\bar{v}_2^2v_1)/\bar{v}_1\bar{v}_2]$$

should equal zero.

$$s_1^2/3-s_2-v_1v_2/3 \text{ should equal zero.}$$

Now

$$v_1v_2=(a+\omega^2b+\omega c)(a+\omega b+\omega^2c)=s_1^2-3s_2.$$

Then we have

$$s_1^2/3-s_2-s_1^2/3+s_2=0;$$

which shows that the triads abc and $ff_1\infty$ are apolar. This gives us

THEOREM 2. *The three vertices of a triangle are a set of points which is apolar to the set consisting of the two Fermat points of the triangle and the point infinity.*

II.

We now take, on a curve K (namely, a biquadratic which cuts itself at right angles), the double point and five other points, and derive the condition

under which the biquadratic having a double point at each and passing through the others, will be rectangular. If we choose K to be a biquadratic cutting itself at right angles at infinity, we know that this is a rectangular hyperbola, and since the other five points must lie on this curve, we see that we may take as their co-ordinates $x_i = m_i \; y_i = 1/m_i \; (i = 1, 2, 3, 4, 5)$. Here x and y are Cartesian co-ordinates, and are not complex.

If we wish to write the biquadratic passing through one of the points, say $x_1 y_1$, which also passes through ∞, we need only write the equation of a circular cubic passing through $x_1 y_1$. This will be of the form

$$(10) \qquad \begin{aligned}&[(x-x_1)^2+(y-y_1)^2]\,[a(x-x_1)+b(y-y_1)]\\ &\quad+c(x-x_1)^2+d(x-x_1)(y-y_1)+e(y-y_1)^2\\ &\quad+f(x-x_1)+g(y-y_1)=0,\end{aligned}$$

(a, b, and c are real, arbitrary constants, and not the a, b, and c of the previous section.) When $x_1 = m_1$ and $y_1 = 1/m_1$ this cuts the hyperbola when

$$(11) \quad \begin{aligned}&[(m-m_1)^2+(1/m-1/m_1)^2]\,[a(m-m_1)+b(1/m-1/m_1)]\\ &\quad+c(m-m_1)^2+d(m-m_1)(1/m-1/m_1)\\ &\quad+e(1/m-1/m_1)^2+f(m-m_1)+g(1/m-1/m_1)=0.\end{aligned}$$

Since this is to have a double point at m_1, $1/m_1$ and cut itself at right angles there, we must have $f = g = 0$ and $e = -c$. Putting these values in equation (11), simplifying, and removing the factor $(m - m_1)^2$, we have

$$(12) \qquad \begin{aligned}&[m^2m_1^2+1]\,[amm_1(m-m_1)-b(m-m_1)]\\ &\qquad\qquad+cm^3m_1^3-dm^2m_1^2-cmm_1=0.\end{aligned}$$

This is obviously a quartic in m, the coefficient of m^4 being am_1^3. We may accordingly write equation (12) as

$$\begin{aligned}&[m^2m_1^2+1]\,[amm_1(m-m_1)-b(m-m_1)]+cm^3m_1^3\\ &\quad-dm^2m_1^2-cmm_1\equiv am_1^3(m-a_1)(m-a_2)(m-a_3)(m-a_4),\end{aligned}$$

pass through the points m_2, $1/m_2$; m_3, $1/m_3$; m_4, $1/m_4$ and m_5, $1/m_5$, where a_1, a_2, a_3 and a_4 are the roots of the quartic. But we wish our curve to Accordingly, we may substitute m_2, m_3, m_4 and m_5 for a_1, a_2, a_3, and a_4 in the last equation, which then becomes

$$(13) \quad \begin{aligned}&[m^2m_1^2+1]\,[amm_1(m-m_1)-b(m-m_1)]+cm^3m_1^3\\ &\quad-dm^2m_1^2-cmm_1\equiv am_1^3(m-m_2)(m-m_3)(m-m_4)(m-m_5).\end{aligned}$$

We are trying to find the conditions under which the biquadratic having a double point at m_1, $1/m_1$ and passing through the four other points m_2, $1/m_2$; m_3, $1/m_3$; m_4, $1/m_4$; m_5, $1/m_5$ and the point at infinity will be

rectangular. This will amount to finding some relation on m_1, m_2, m_3, m_4 and m_5. In (13) we have a relation involving all five m's, but it also involves the arbitrary constants a, b, c and d. We must get rid of these constants in some way. To do this, we take equation (13) and let $m=m_1$, $mm_1=i$ and $mm_1=-i$ successively, thus obtaining the three equations

$$(14) \qquad c(m_1^6-m_1^2)-dm_1^4=am_1^3(m_1^4-m_1^3\sigma_1+m_1^2\sigma_2-m_1\sigma_3+\sigma_4),$$

$$(15) \qquad -2ci+d=(a/m_1)(i^4-i^3m_1\sigma_1+i^2m_1^2\sigma_2-im_1^3\sigma_3+m_1^4\sigma_4),$$

$$(16) \qquad 2ci+d=(a/m_1)(i^4+i^3m_1\sigma_1+i^2m_1^2\sigma_2+im_1^3\sigma_3+m_1^4\sigma_4),$$

where σ_1, σ_2, σ_3 and σ_4 are the four symmetric functions of m_2, m_3, m_4 and m_5.

Solving the simultaneous equations (15) and (16) we obtain $c=a/2\,(m_1^2\,\sigma_3-\sigma_1)$ and $d=a/m_1(1-m_1^2\,\sigma_2+m_1^4\,\sigma_4)$. Substituting these back in equation (14), we find that each term contains a as a factor, so that this may be cancelled out, leaving an expression which involves only m_1 and the symmetric functions of the other four m's. This expression is, after some reductions have been made,

$$(17) \qquad \sigma_1+m_1^2\,\sigma_3-2m_1\,\sigma_4-2m_1=0.$$

This equation involves m_1 in a way different from that in which the remaining four m's are involved. This is because we started out by letting our biquadratic have a double point at m_1, $1/m_1$. But we wish to treat all the m's alike; i. e., we wish to find the condition under which the biquadratic having a double point at *any* m and passing through the others will be rectangular. So we must find a relation which will treat all the m's alike. To do this, we introduce the symmetric functions of m_1, m_2, m_3, m_4 and m_5, which are denoted by s_1, s_2, s_3, s_4 and s_5, and which are related to the σ's as follows:

$$s_1=m_1+\sigma_1, \qquad s_3=m_1\sigma_2+\sigma_3, \qquad s_5=m_1\sigma_4.$$
$$s_2=m_1\sigma_1+\sigma_2, \qquad s_4=m_1\sigma_3+\sigma_4,$$

Also, we have

$$\sigma_1=s_1-m_1, \qquad \sigma_3=(s_4m_1-s_5)/m_1^2, \qquad \sigma_4=s_5/m_1.$$

Putting these into equation (17) and simplifying, we have:

$$(18) \qquad m_1(s_4-3)+s_1-3s_5=0.$$

Equating coefficients of like powers of m_1 gives

$$(19) \qquad s_4=3, \qquad s_1=3s_5.$$

Equations (19) give the condition which we sought; it proves to be two

relations on m_1, m_2, m_3, m_4 and m_5; if we desire to have just one condition, we may effect this by substituting for the s's their values in terms of the σ's; this gives: $m_1\,\sigma_3+\sigma_4=3$ and $m_1+\sigma_1=3m_1\,\sigma_4$. Elimination of m_1 from these two equations gives

$$\sigma_1\sigma_3-10\sigma_4+3\sigma_4^2+3=0. \tag{20}$$

This is a relation on m_2, m_3, m_4 and m_5, but it will hold for any four of the m's. For instance, we might have started working with a biquadratic having a double point at m_2, and then we should have arrived at a relation (20) involving m_1, m_3, m_4 and m_5. So (20) holds for any four of the m's.

Equation (20) is fundamental in this work, so we state

THEOREM 3. *If we take, on a biquadratic which cuts itself at right angles at infinity, the point infinity and five other points whose co-ordinates are* m_1, $1/m_1$; m_2, $1/m_2$; m_3, $1/m_3$; m_4, $1/m_4$ *and* m_5, $1/m_5$, *then the condition that the biquadratic having a double point at* m_1, $1/m_1$ *and passing through the other five points shall be rectangular is:* $\sigma_1\sigma_3-10\sigma_4+3\sigma_4^2+3=0$, *where the* σ*'s are the symmetric functions of* m_2, m_3, m_4 *and* m_5.

III.

In order to get a geometrical interpretation of the condition given in equation (20), we investigate the connection between our six points and Neuberg's cubic curve for three points.

This curve is defined as follows: starting out with three points $x_1\,y_1$, $x_2\,y_2$, $x_3\,y_3$, Neuberg's cubic is the locus of points $x_4\,y_4$ satisfying the condition.*

$$\Delta'\equiv\begin{vmatrix}\lambda_{23}\lambda_{14} & \lambda_{23}+\lambda_{14} & 1\\ \lambda_{31}\lambda_{24} & \lambda_{31}+\lambda_{24} & 1\\ \lambda_{12}\lambda_{34} & \lambda_{12}+\lambda_{34} & 1\end{vmatrix}=0,$$

where λ_{23} is the squared distance between the points x_2y_2 and x_3y_3, and similarly for the other λ's. This curve passes through the points x_1y_1, x_2y_2 and x_3y_3. In other words, if four points lie on a Neuberg cubic, Δ' vanishes for them.

If we subject the four points involved in Δ' to an inversion with center at x_5y_5 and radius equal to k, we get

$$\Delta'=k^{12}I_2/\lambda_{15}^2\lambda_{25}^2\lambda_{35}^2\lambda_{45}^2,$$

* See the note by Professor Morley, referred to above.

where

$$I_2 \equiv \begin{vmatrix} \lambda_{23}\lambda_{14} & \lambda_{23}\lambda_{15}\lambda_{45}+\lambda_{14}\lambda_{25}\lambda_{35} & 1 \\ \lambda_{31}\lambda_{24} & \lambda_{31}\lambda_{25}\lambda_{45}+\lambda_{24}\lambda_{35}\lambda_{15} & 1 \\ \lambda_{12}\lambda_{34} & \lambda_{12}\lambda_{35}\lambda_{45}+\lambda_{34}\lambda_{15}\lambda_{25} & 1 \end{vmatrix} = 0.$$

If $\Delta'=0$ is the original Neuberg cubic, then $I_2=0$ is a new curve, which is the locus of a point x_4y_4, the other four points being given. The curve whose equation is $I_2=0$ is a biquadratic; it is a covariant biquadratic curve of the four points x_1, y_1; x_2, y_2; x_3, y_3 and x_5, y_5. This curve, which may be denoted by C_2, is the "inverted Neuberg cubic with respect to one of the points (say x_5, y_5) of any three of the remaining four (say x_1, y_1; x_2, y_2 and x_3, y_3). The fifth point x_4, y_4 is the variable point which traces out the locus." * If five points lie on an inverted Neuberg cubic, I_2 must vanish for them.

It should be noticed that while the Neuberg cubic seems to involve only four points, it really involves five, since it is a biquadratic on infinity. So it involves the four points which are explicitly contained in it, plus the point infinity. The inverted Neuberg cubic, or better, covariant biquadratic C_2 involves five points explicitly, since, on inversion, infinity goes into $x_5\,y_5$, and this appears in the equation.

After this digression on the Neuberg cubic and the covariant biquadratic C_2, we discuss the connection between these curves and our set of six points.

If we wish four of our points, say $m_2, 1/m_2$; $m_3, 1/m_3$; $m_4, 1/m_4$, and $m_5, 1/m_5$ to lie on a Neuberg cubic, we must have

$$\begin{vmatrix} \left[(m_3-m_4)^2+\frac{(m_4-m_3)^2}{m_4^2m_3^2}\right]\left[(m_2-m_5)^2+\frac{(m_5-m_2)^2}{m_5^2m_2^2}\right] & (m_3-m_4)^2+\frac{(m_4-m_3)^2}{m_4^2m_3^2}+(m_2-m_5)^2+\frac{(m_5-m_2)^2}{m_5^2m_2^2} & 1 \\ \left[(m_4-m_2)^2+\frac{(m_2-m_4)^2}{m_2^2m_4^2}\right]\left[(m_3-m_5)^2+\frac{(m_5-m_3)^2}{m_3^2m_5^2}\right] & (m_4-m_2)^2+\frac{(m_2-m_4)^2}{m_2^2m_4^2}+(m_3-m_5)^2+\frac{(m_5-m_3)^2}{m_3^2m_5^2} & 1 \\ \left[(m_2-m_3)^2+\frac{(m_3-m_2)^2}{m_2^2m_3^2}\right]\left[(m_4-m_5)^2+\frac{(m_5-m_4)^2}{m_4^2m_5^2}\right] & (m_2-m_3)^2+\frac{(m_3-m_2)^2}{m_2^2m_3^2}+(m_4-m_5)^2+\frac{(m_5-m_4)^2}{m_4^2m_5^2} & 1 \end{vmatrix}=$$

The problem of evaluating this determinant is quite interesting in itself, but since it involves many detailed operations which would be tedious for the

* Morley, *loc. cit.*, page 408.

reader, and which are unimportant for the purpose of this paper, we may omit them, and let it suffice to say that the determinant may finally be reduced to

$$(m_2m_3m_4m_5-1)^2\begin{vmatrix}(A-B)^2 & C & 1\\ (A-C)^2 & B & 1\\ (B-C)^2 & A & 1\end{vmatrix}+\begin{vmatrix}C^2(A-B)^2 & C & 1\\ B^2(A-C)^2 & B & 1\\ A^2(B-C)^2 & A & 1\end{vmatrix}=0,$$

where

$$A=m_2m_3+m_4m_5,\qquad B=m_2m_4+m_3m_5,\qquad C=m_2m_5+m_3m_4.$$

This gives

$$(\sigma_4-1)^2[-3(A-B)(B-C)(C-A)]$$
$$-(AB+AC+BC)(A-B)(B-C)(C-A)=0.$$

Then either

$$3(\sigma_4-1)^2+(AB+AC+BC)=0 \text{ or } (A-B)(B-C)(C-A)=0.$$

Making the second expression equal zero makes one of its factors zero. If, say, $A-B=0$, then $A=B$ and $m_2m_3+m_4m_5=m_2m_4+m_3m_5$. This makes either $m_5=m_2$ or $m_4=m_3$. We assumed our six points to be distinct, so this relation cannot hold. Therefore, we have

$$3(\sigma_4-1)^2+(AB+AC+BC)=0,\qquad 3(\sigma_4-1)^2+\sigma_1\sigma_3-4\sigma_4=0,$$
$$3\sigma_4^2-6\sigma_4-4\sigma_4+\sigma_1\sigma_3+3=0,\qquad \sigma_1\sigma_3-10\sigma_4+3\sigma_4^2+3=0.$$

Thus the condition that four of the points lie on a Neuberg cubic is exactly the same as the condition that the biquadratic having a double point at one of the six points and passing through the others shall be rectangular, since in each case the condition is

$$\sigma_1\sigma_3-10\sigma_4+3\sigma_4^2+3=0.$$

Since this holds for *any four* of our five points, we know that *any four* of them lie on a Neuberg cubic, or Δ' vanishes for any four of our five finite points. However, this really involves not four but *five* points. For when we are given a set of n points in the inversive plane, we have, in reality, $n+1$ points, for the point infinity may always be regarded as an additional point of the set. So this condition involves, in reality, five points of our set, namely, four of the five finite points and infinity. If we do not like to bother with the point infinity, we can invert with respect to a point a_6. So that now, instead of dealing with a set of points m_1, $1/m_1$; m_2, $1/m_2$; m_3, $1/m_3$; m_4, $1/m_4$; m_5, $1/m_5$ and infinity, we are dealing with a set of points a_1, a_2, a_3, a_4, a_5 and a_6, where $a_i(i=1\cdots5)$ is the inverse of the point m_i, $1/m_i$

and a_6 is the inverse of infinity. The old condition for rectangularity will invert into a new condition, and this will be the condition that the biquadratic having a double point at one of the six (finite) points $a_1, \cdots, a_6$ and passing through the others shall be rectangular. The old condition involved four points, and said that any four of the five finite points in the old set lay on a Neuberg cubic; the new condition involves five points and says that any *five* of the points $a_1, \cdots, a_6$ lie on a covariant curve C_2 (since a Neuberg cubic inverts into a C_2), or better, that I_2 vanishes for any five of the six points.

Now, suppose we investigate more closely our set of points $a_1, \cdots, a_6$. Four of these, say a_2, a_3, a_4 and a_5 will determine a covariant curve C_2 which passes through a_2, a_3, a_4 and a_5. *Any point on this curve,* together with the points a_2, a_3, a_4 and a_5, will satisfy the condition $I_2 = 0$. We have already seen that any five of our set of points $a_1, \cdots, a_6$ satisfy the condition $I_2 = 0$. That is, I_2 will be zero for the points a_1, a_2, a_3, a_4 and a_5 and also for the set of points a_6, a_2, a_3, a_4 and a_5. From this it is evident that a_1 and a_6 lie on the curve determined by a_2, a_3, a_4 and a_5; that is, the C_2 determined by a_2, a_3, a_4 and a_5 passes through a_1 and a_6. In the same way we see that the C_2 determined by *any four* of the points is on the other two.

We may also put it as follows:

Choose a set of five of the points, say a_2, a_3, a_4, a_5 and a_6. Then

a_2, a_3, a_4, a_5 determine a C_2 which passes through a_1 and a_6,
a_2, a_3, a_4, a_6 determine a C_2 which passes through a_1 and a_5,
a_2, a_3, a_5, a_6 determine a C_2 which passes through a_1 and a_4,
a_2, a_4, a_5, a_6 determine a C_2 which passes through a_1 and a_3,
a_3, a_4, a_5, a_6 determine a C_2 which passes through a_1 and a_2.

It is to be noticed that all these C_2*'s* pass through a_1. Thus we may say: If we choose a set of five of the six points, then, by choosing sets of four points from this set of five points, we determine five C_2*'s*. All five of these C_2*'s* then pass through the sixth point.

The results of the last three paragraphs are the most important ones arrived at in this paper. Accordingly, we re-state them, for better emphasis, as

THEOREM 4. *If we have a set of six points, which are such that the biquadratic having a double point at each, and passing through the others is rectangular, then the invariant I_2 vanishes for any five of them. The covariant curve C_2 determined by any four is on the other two. Moreover, if we choose a set of five of the six points, and set up the five covariant curves C_2 which are determined by sets of four points chosen from this set of five points, then all these five C_2's pass through the sixth point.*

IV.

It is proper to regard the six points and associated biquadratics, which we have studied so far, as the fundamental points and curves of a Geiser involution in the inversive plane.

In the inversive plane, circles play the rôle which lines play in the projective plane. A Cremona transformation in this plane is one which transforms circles into curves which behave like circles; i. e., into curves C_n* which are rational, which form a triply infinite system, and which are such that any two of them have as many *free* points of intersection as do two circles.

Professor Morley, starting from this definition,† has determined that there are four symmetric Cremona transformations in the inversive plane, and has given some of their properties. The transformations are:

A. One which transforms circles into circles,
B. One which transforms circles into C_3's *with four double points,*
C. One which transforms circles into C_7's *with six four-fold points,*
D. One which transforms circles into C_{15}'s *with seven eight-fold points.*

These transformations are all involutory.

Type C involves six points, and is therefore the one to which we shall confine our attention. This transformation sends a circle into a C_7 with six four-fold points. These six points are the fundamental points of the transformation. What are the fundamental curves? They are, of course, the transforms of the fundamental points. The transformation sends the C_7 on which a given fundamental point F is situated into a circle. Since the point F is of multiplicity four, a point P, in describing C_7, passes through F four times. The transform of P, called P', describes a circle, and this circle must meet the transform of F, which is a curve C_n, just four times. A circle and a C_n meet $2n$ times; therefore, $2n = 4$, $n = 2$, and the fundamental curve is a biquadratic, or C_2. As a matter of fact, it can be shown that the biquadratic corresponding to a given fundamental point has a double point at that point, and passes through the remaining five fundamental points. Thus our fundamental system consists of six points and six biquadratics, each biquadratic having a double point at one of the six points, and passing through the others.

To a general circle corresponds then a C_7 with four-fold points a_i

* See explanatory note at end of paper.

† Unpublished report in mathematics seminary, Johns Hopkins University, December, 1926. The question as to whether a transformation so defined is one-to-one is not of interest here; it was, however, discussed by Professor Morley in the report mentioned here.

$(i=1,\cdots,6)$. When the circle is on a_1, the corresponding fundamental biquadratic A_1 is a factor of C_7; the circle then corresponds to the other factor, a C_5 with a double point at a_1 and triple points at the other a's.

When the circle is also on a_2, then A_2 factors out from C_5, and what corresponds to the circle is a C_3, with a_1, a_2 as simple points and $a_3\,a_4\,a_5\,a_6$ double points.

Finally, to the circle on a_1, a_2, a_3 corresponds the circle on a_4, a_5, a_6.

To what Cremona transformation in the projective plane does this transformation correspond? In the projective plane, a Cremona transformation transforms lines into curves which behave like lines. So we must see what the effect of our transformation on a line will be. In the inversive plane, a line is simply a circle on infinity. Let us put one of the fundamental points of our transformation at infinity. Then, by (1) above, the transform of a line will be a C_5 with infinity as a double point, and five triple points. But a C_5 with infinity as a double point is, projectively, an octavic with I and J as triple points. Thus we see that our transformation corresponds to one in the projective plane which sends a line into an octavic with seven triple points. This is the well-known Geiser involution, and by analogy we call Type C the Geiser involution in the inversive plane. We can now restate our main theorem thus:

There is a Geiser involution for which the six biquadratics are all rectangular. This requires three conditions on the six points.

Note on Curves in the Inversive Plane.

For the sake of those readers who are not familiar with the handling of curves in the inversive plane, I append the following note:

In this plane we use as co-ordinates $z=x+iy$ and $\bar{z}=x-iy$ (x and y being the ordinary Cartesian co-ordinates of a point.) These are variously called minimal, circular or absolute co-ordinates. A curve is given by an expression in z and $\bar{z}$, equated to zero. It is generally written in the form of a self-conjugate matrix such as

$$\begin{array}{c|cccccccc}
 & 1 & z & z^2 & z^3 & z^4 & \cdots & z^n \\
\hline
1 & P_0 & a_1 & a_2 & a_3 & a_4 & \cdots & a_n \\
\bar{z} & \bar{a}_1 & P_1 & b_1 & b_2 & b_3 & \cdots & b_{n-1} \\
\bar{z}^2 & \bar{a}_2 & \bar{b}_1 & P_2 & c_1 & c_2 & \cdots & c_{n-2} \\
\bar{z}^3 & \bar{a}_3 & \bar{b}_2 & \bar{c}_1 & P_3 & d_1 & \cdots & d_{n-3} \\
\bar{z}^4 & \bar{a}_4 & \bar{b}_3 & \bar{c}_2 & \bar{d}_1 & P_4 & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \ddots & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \ddots & \\
\cdot & \cdot & \cdot & \cdot & \cdot & & & \ddots \\
\bar{z}^n & \bar{a}_n & \bar{b}_{n-1} & \bar{c}_{n-2} & \bar{d}_{n-3} & & & P_n
\end{array} = 0,$$

where the p's are real and the other numbers are complex quantities and their conjugates. This is read by multiplying every element by the quantities heading the row and column in which the element stands and adding all such products together. Thus, the curve written above would start off as

$$P_0 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4 + \cdots + a_n z^n$$
$$+ \bar{a}_1 \bar{z} + P_1 z\bar{z} + b_1 z^2 \bar{z} + b_2 z^3 \bar{z} + \cdots = 0.$$

A curve is always named after the highest powers of z and $\bar{z}$ occurring in it. For instance, if the term of highest power is $p_n z^n \bar{z}^n$ the curve is called a bi-n-ic, and is denoted by C_n. It is easily seen that this curve is, projectively, a $2n$-*ic*, with an n-fold point at I and an n-fold point at J, since if we express $p_n z^n \bar{z}^n$ in Cartesian co-ordinates it becomes $p_n (x^2 + y^2)^n$ and this is the leading term of the curve when written in the ordinary way. The simplest curve is the C_1

$$\begin{array}{c|cc} & 1 & z \\ \hline 1 & P_0 & a_1 \\ \bar{z} & \bar{a}_1 & P_1 \end{array} = 0,$$

or $p_0 + a_1 z + \bar{a}_1 \bar{z} + p_1 z\bar{z} = 0$, which is a circle, since it is a conic on I and J. The next simplest is the C_2 or biquadratic

$$\begin{array}{c|ccc} & 1 & z & z^2 \\ \hline 1 & P_0 & a_1 & a_2 \\ \bar{z} & \bar{a}_1 & P_1 & b_1 \\ \bar{z}^2 & \bar{a}_2 & \bar{b}_1 & P_2 \end{array} = 0,$$

or

$$P_0 + a_1 z + \bar{a}_1 \bar{z} + a_2 z^2 + \bar{a}_2 \bar{z}^2 + P_1 z\bar{z} + b_1 z^2 \bar{z} + \bar{b}_1 \bar{z}^2 z + P_2 z^2 \bar{z}^2 = 0$$

which is projectively a quartic with I and J as double points.

If, in a C_n, the coefficient of $z^n \bar{z}^n$ is zero, the C_n passes through infinity. If, in addition, the coefficients of $z^n\ \bar{z}^{n-1}$ and $\bar{z}^n\ \bar{z}^{n-1}$ are both zero, the curve has a double point at infinity. Thus, writing the curve in matrix form, we see that by striking off the element in the lower right-hand corner, we make the curve pass through infinity. This causes the curve to be, projectively speaking, a $(2n-1)$ic, with I and J as $(n-1)$ fold points, for our highest term will be of the form

$$\begin{aligned} kz^n \bar{z}^{n-1} + \bar{k}\bar{z}^n z^{n-1} &= z^{n-1}\bar{z}^{n-1}[z(k_1 + ik_2) + \bar{z}\,(k_1 - ik_2)] \\ &= z^{n-1}\bar{z}^{n-1}[(k_1 + ik_2)(x + iy) + (k_1 - ik_2)(x - iy)] \\ &= (x^2 + y^2)^{n-1}[2k_1 x - 2k_2 y]. \end{aligned}$$

Taking off in addition the next two elements in the lower right-hand corner makes infinity a double point, and makes the curve equivalent to a curve in the projective plane of order $2n - 2$, with $(n - 2)$-fold points at I and J. We see immediately that a circle on infinity is a straight line, while a C_2 on infinity is a circular cubic. A C_2 with a double point at infinity is a conic.

In general, we may say that starting in the lower right-hand corner of the matrix and striking off p diagonals in succession makes our C_n have a p-fold point at infinity; it is a curve in the projective plane of order $(2n - p)$ with $(n - p)$ fold points at I and J. Given a curve in the inversive plane whose highest terms are of the form $kz^p\bar{z}^q$ and $\bar{k}\bar{z}^p z^q$, where $p \neq q$, then the curve is a C_n, where n is the greater of the two integers p and q, which passes through infinity $p - q$ or $q - p$ times. It is, in the projective plane a curve of order $p + q$ which passes through I and J each either q or p times. Thus, a curve whose highest terms are of the form $z^2\bar{z}$ and $\bar{z}^2 z$ is a C_2 with infinity as a simple point. It is projectively a curve of the third degree which passes simply through I and J; i. e., a circular cubic.

A C_n with an n-fold point at infinity is a curve in the projective plane of order n, not passing through I and J. Accordingly we may say that an ordinary curve of order n in the projective plane is, in the inversive plane, a C_n with an n-fold point at infinity. A curve of order n in the projective plane which passes through I and J each m times becomes, in the inversive plane, a C_{n-m} with infinity as $(n - 2m)$fold point, whose highest terms are of the form $z^{n-m}\bar{z}^m$ and $z^m\bar{z}^{n-m}$. Thus, a curve of order nine in the projective plane, which passes through I and J each three times, becomes a C_6 with infinity as a triple point, whose highest terms are of the form $z^6\bar{z}^3$ and $\bar{z}^6 z^3$.

One more point concerning these curves must be called to the attention of the reader: A C_n and a C_m have in common $2mn$ points.* For the C_n, being a $2n$-ic with n-fold points at I and J, and the C_m being a $2m$-ic with m-fold points at I and J, intersect in $4mn$ points; but the intersections at I and J use up $2mn$ of these, so that, as a consequence, we may say that they have in common only $4mn - 2mn = 2mn$ points.

* The common points are either actual intersections or common image-pairs.

Concerning the Rational Curves $R_3^5(II)$ and R_2^4.*

By H. E. Arnold.

1. The rational space quintic curve, $R_3{}^5$, is given parametrically by four binary quintics

$$(1) \qquad R_3{}^5: \qquad x_i = (\alpha_i t)^5, \qquad (i = 0, 1, 2, 3).$$

The pencil of apolar quintics

$$(2) \qquad \rho(bt)^5 + \sigma(ct)^5$$

has been called the fundamental involution, and the necessary and sufficient condition that a binary quintic represent a plane section of the curve is that it be apolar to the fundamental involution.

In general there is a unique quartic apolar to the fundamental involution, and therefore a rational space quintic possesses a unique quadrisecant line. But if the four conditions imposed on the binary quartic by making it apolar to the quintics (2) are not independent, then there is a pencil of quartics apolar to the fundamental involution. This means that the curve possesses an infinity of quadrisecant lines, and it is said to be of the second species.† We shall designate this type of curve by the symbol $R_3{}^5(II)$, and shall call the pencil of apolar quartics,

$$(3) \qquad \lambda(\beta t)^4 + \mu(\gamma t)^4,$$

the "auxiliary involution." The curve lies on a quadric and the ∞' quadrisecant lines form one set of rulings.

The rational plane quartic curve is represented by three binary quartics

$$(4) \qquad R_2{}^4: \qquad x_i = (\alpha_i t)^4, \qquad (i = 0, 1, 2),$$

and the fundamental pencil is the involution of quartics apolar to (4). The form (3) will be considered also as the fundamental pencil of $R_2{}^4$.

The two curves $R_2{}^4$ and $R_3{}^5(II)$ can therefore be regarded as being determined, to within projections, by a pencil of binary quartics, or by means of the unique binary sextic, Ψ, apolar to such a pencil.

Setting

$$(5) \qquad p_{ij} = \begin{vmatrix} \beta_i & \beta_j \\ \gamma_i & \gamma_j \end{vmatrix},$$

* Presented to the American Mathematical Society, October 27, 1928.

† Bertini, "Sulle Curve Gobbe Razionali Del. 5. Ordine," *Collectanea Mathematica in mem. D. Chelini.*

the combinants * of (3) are expressed in terms of these determinants. The invariants for various special types of the curve R_2^4 have been found † and in the following table we give the relationship of some of these forms to the quintic $R_3^5(II)$.

Conditions on Pencil of Quartics	Special Quadrisecant lines of $R_3^5(II)$	Quartic R_2^4 possesses
$I_4-(I_2'-36I_2)^2=0$	1 bitangent	Degenerate flex conic
$I_4-(I_2'-36I_2)^2=0$ $I_6-16I_2^3=0$	2 bitangents	Biflecnode
$I_4-(I_2'-36I_2)^2=0$ $I_6-16I_2^3=0$ $(qt)^2\equiv 0$	3 bitangents	3 biflecnodes
$I_6=0$	1 linear inflexion	1 cusp
$I_6=0$ $[(I_2'-4I_2)^2-I_4]^2$ $+128[I_2I_2'I_4-I_2(I_2'-4I_2)^3+28I_2^2I_4]=0$	2 linear inflexions	2 cusps
$I_6=0$ $[(I_2'-4I_2)^2-I_4]^2$ $+128[I_2I_2'I_4-I_2(I_2'-4I_2)^3+28I_2^2I_4]=0$ $I_4-64I_2(I_2'-4I_2)-(I_2'-4I_2)^2=0$	3 linear inflexions	3 cusps
Apolar sextic Ψ has triple root	1 four-point contact tangent	R_2^4 reduces to rational cubic ‡
Apolar sextic Ψ is cube of a quadratic	2 four-point contact tangents	R_2^4 reduces to conic

* The relations between the Morley and the Salmon combinants (Salmon, *Modern Higher Algebra*, 4th Edition, pp. 219-221) are: $I_2=-B$; $I_2'=12B-A$; $I_4=R$; $I_6=D$; $(qt)^2=P$; $(st)^6=J$.

† Rowe, "Covariants and Invariants of the Rational Plane Quartic," *Transactions of the American Mathematical Society*, Vol. 12, p. 304; Winger, "Self-Projective Rational Curves of the Fourth and Fifth Orders," *American Journal of Mathematics*, Vol. 36, p. 59.

‡ Thomsen, "The Osculants of Plane Rational Quartic Curves," *American Journal of Mathematics*, Vol. 32, p. 207.

2. The curve $R_3^5(II)$ for which I_2 vanishes is particularly interesting. The vanishing of I_2 is the condition that the pencil (3) be the first polars of a binary quintic. For the plane rational quartic $I_2=0$ is the triple point condition, and it is known that the canonizant of the quintic form gives the parameters of the triple point.

Let us take as the base quintic form

$$(6)\qquad (at)^5=a_0t^5+5a_1t^4\tau+10a_2t^3\tau^2+10a_3t^2\tau^3+5a_4t\tau^4+a_5\tau^5.$$

The fundamental involution is readily obtained from the auxiliary involution, and we note that the five points of the curve which are given by the roots of (6) lie on a plane, which we shall call the plane π. Let us designate by $(at_1)(at)^4$ and $(at_2)(at)^4$, respectively, the first polars of (t_1,τ_1) and (t_2,τ_2) with respect to (6). It can easily be verified that

$$(7)\qquad (t\tau_2-t_2\tau)(at_1)(at)^4-(t\tau_1-t_1\tau)(at_2)(at)^4=(t_1\tau_2-t_2\tau_1)(at)^5.$$

But

$$(8)\qquad (t\tau_2-t_2\tau)(at_1)(at)^4$$

and

$$(9)\qquad (t\tau_1-t_1\tau)(at_2)(at)^4$$

are plane sections of the curve. Moreover, when the quintics giving three plane sections are linearly related the planes are in a pencil, and hence the planes (8) and (9) intersect on π. Thus if P_1 and P_2 are any two points of the curve, and q_1 and q_2 their polar quadrisecants, respectively, we have the planes P_1q_2 and P_2q_1 intersecting on the plane π.

If we let $R_3^5(II)$ be given by

$$(10)\qquad R_3^5(II)\quad \begin{cases} x_0=t[\lambda_1(\beta t)^4+\mu_1(\gamma t)^4] \\ x_1=\tau[\lambda_1(\beta t)^4+\mu_1(\gamma t)^4] \\ x_2=t[\lambda_2(\beta t)^4+\mu_2(\gamma t)^4] \\ x_3=\tau[\lambda_2(\beta t)^4+\mu_2(\gamma t)^4], \end{cases}$$

where $\lambda_1(\beta t)^4+\mu_1(\gamma t)^4$ and $\lambda_2(\beta t)^4+\mu_2(\gamma t)^4$ are any two members of the auxiliary involution, then the curve lies on the quadric

$$(11)\qquad x_0x_3-x_1x_2=0,$$

and the equation of the plane π is

$$(12)\qquad \mu_2x_0-\lambda_2x_1-\mu_1x_2+\lambda_1x_3=0.$$

The equations of the plane sections (8) and (9) are given by

$$(13)\qquad (t_1\mu_2-\tau_1\lambda_2)(\tau_2x_0-t_2x_1)+(\tau_1\lambda_1-t_1\mu_1)(\tau_2x_2-t_2x_3)=0$$

and

$$(14)\qquad (t_2\mu_2-\tau_2\lambda_2)(\tau_1x_0-t_1x_1)+(\tau_2\lambda_1-t_2\mu_1)(\tau_1x_2-t_1x_3)=0,$$

respectively.

If C be the conic in which π cuts the quadric (11), and if S_1 and S_2 be the points in which C is met by the quadrisecants q_1 and q_2, respectively, then as P_2 moves along the curve while P_1 remains fixed, the planes (13) and (14) give rise to a pencil of lines through S_1. As S_2 is brought into coincidence with S_1 the line S_1S_2 becomes tangent to C, and we have the theorem: *If I_2 vanishes the plane through a point P_1 of the curve and its polar quadrisecant q_1 cuts the plane π in a line which is tangent to the conic C.* A simple geometric method for finding the pole of a given quadrisecant is thus afforded.

But the tangent line to C at S_1 lies in the plane of the two generators through that point, and hence *the conic C may be further characterized as being the locus of points through which pass a unisecant and a quadrisecant generator such that the point on the unisecant is the pole, with respect to $(at)^5$, of the quartic giving the quadrisecant.* The curve $R_3^5(II)$ itself may then be considered as being specialized in this way when the invariant I_2 vanishes.

If we let $t=0$ be one of the roots of (6), then the polar of $(0,1)$ gives the quadrisecant through this point. But $t=0$ is then also a root of the polar quadrisecant, and hence one of the four points has moved up to the conic C. This gives the theorem: *At the five points whose parameters are the roots of the base quintic (6), the tangent line to the curve, the tangent to the conic, and the quadrisecant through the point considered are coplanar.*

3. In the theory of the rational plane quartic the undulation invariant I_4 was obtained as the eliminant of the two forms $(\beta t)^4$ and $(\gamma t)^4$ of the fundamental pencil. If we impose the condition that the I_4 of the auxiliary involution of $R_3^5(II)$ vanish, then each member of the system of quadrisecants would have a root in common. Since the quadrisecants form one system of generators of the quadric on which the curve lies, that surface would have to be a quadric cone. But then a plane through two of the elements of the cone would cut the quintic in seven points. Similarly, if $(\beta t)^4$ and $(\gamma t)^4$ had two roots in common a plane through two elements would meet $R_3^5(II)$ in six points. If, however, the two quartics have three roots in common the curve has a triple point at the vertex of the cone.* We

* Mueller, Dissertation, *Die Rationale Kurve Fünfter Ordnung im Fünf-, Vier-, Drei- und Zweidimensionalen Raum.* Leipzig (1910), p. 78.

conclude, therefore, that I_4 *cannot vanish for any* $R_3^5(II)$ *which does not have a triple point.*

The $R_3^5(II)$ with a triple point corresponds, then, to the R_2^4 with three undulations. As in the case of such a plane quartic the sextic Ψ for $R_3^5(II)$ is not a unique form, but any member of a net of sextics will serve to recover the curve.*

4. The $R_3^5(II)$ which possesses two four-point contact tangents is interesting because it sets up a null system. Taking the equations of such a curve as given by Colpitts,†

$$(15) \qquad R_3^5(II): \quad x_0 = t^5, \quad x_1 = t\tau^4, \quad x_2 = t^4\tau, \quad x_3 = \tau^5,$$

the equations in planes are found to be

$$(16) \qquad \xi_0 = 3\tau^5, \quad \xi_1 = 5t^4\tau, \quad \xi_2 = -5t\tau^4, \quad \xi_3 = -3t^5.$$

Hence it can be shown without difficulty that the five planes which osculate the curve at the points cut out by the plane $(\xi x) = 0$ meet in a point which itself lies on (ξx); dually, a point in space gives rise to a plane through itself.

5. A pencil of binary quartics has also been studied from the point of view of a pencil of line sections of R_2^4. In what follows we will make use of this interpretation.

Consider an R_2^4 with distinct nodes B, C, D, and the ∞^5 cubics C_2^3 which pass through B, C, D, and any fourth point A not on R_2^4. Each such cubic C_2^3 cuts R_2^4 in six more points, and we have therefore ∞^5 binary sextics. Such a system of cubics is determined by six linearly independent members, and likewise the involution of sextics is defined by the six corresponding sextics. But there is a unique sextic $f^6 = (dt)^6$ apolar to the six independent sextics. Any sextic apolar to f^6 is a member of the involution, and one cubic C_2^3 passes through the six points corresponding to the given sextic. The roots of f^6 give the contacts of the six cubics C_2^3 which have six-point contact with R_2^4.

Now any line through A, together with a conic on B, C, and D is a degenerate cubic C_2^3. This means that the four points on a line through A, with any two arbitrary points of the curve R_2^4, give a sextic apolar to f^6. In other words, the quartics obtained from the pencil of lines through A are apolar to f^6. Hence the theorem: *The roots of the unique sextic apolar*

---

* Neelley, "Compound Singularities of the Rational Plane Quartic Curve," *American Journal of Mathematics*, Vol. 49, p. 391.

† Colpitts, "On Twisted Quintic Curves," *American Journal of Mathematics*, Vol. 29, p. 336.

to the pencil of quartics obtained from the line sections through a fixed point not on R_2^4 are the parameters of the contacts of the six cubics C_2^3 which pass through the nodes of R_2^4 and the fixed point and have six-fold contact with the curve.

Furthermore, when we examine the degenerate cubics C_2^3 which pass through one and two fixed points, respectively, of R_2^4, we obtain the theorem:

The four points on a line through A are such that any three of them form the canonizant of the first polar of the fourth with respect of f^6; the second polars of f^6 with respect to two points on a line through A give harmonic sets.

We note particularly the case where the point E on the quartic, taken with the points A, B, C, D, determines a conic, which in turn cuts out another point of the curve. Calling this sixth point F, we see that the net of line sections of R_2^4 gives rise to quartics apolar to the second polars $(dt_E)(dt_F)(dt)^4$, where E and F are in an involution $I_{1,1}^2$ formed by the conics on A, B, C, D. Thus the quartics of the net of line sections are not apolar to a net of quartics,—the second polars of f^6—as might seem to be the case, but to the pencil of quartics obtained from the quadratic involution. We may take as the base quartics of the pencil the pure second polars of f^6 with respect to the two points E_1 and E_2 at which conics on A, B, C, D have contact with R_2^4. Designating this pencil by

$$(17) \qquad \lambda(dt_{E_1})^2(dt)^4 + \mu(dt_{E_2})^2(dt)^4,$$

and noting that the net of line sections of R_2^4 is apolar to (17), we see that (17) is the fundamental pencil of the curve. The net of second polars of f^6 therefore contains the fundamental pencil. We can thus say that

The ∞^2 points of the plane give rise to a net of binary sextics, f^6, each including among its second polars the fundamental pencil of the quartic R_2^4.

On Varieties of Three Dimensions with Six Right Lines through Each Point.

By Charles H. Sisam.

Among surfaces, it is a well known and quite useful theorem that if, through each point of the surface, there pass just two distinct right lines lying on the surface, then the surface is a quadric. It is the purpose of this paper to prove an analogous theorem for varieties of three dimensions; namely, *if through each point of the variety there pass just six right lines which lie on the variety, then the variety is a cubic hypersurface in four dimensions.*

Denote by V_3 such a variety of three dimensions. Since each rectilinear generator at a point P of V_3 is a fourpoint tangent to V_3 at P, the variety has, at each of its points, six fourpoint tangents and lies in a space of four dimensions.* No three of the generators through a generic point can be coplanar; otherwise the quadric cone of three point tangents to V_3 would be composite and the six generators through that point would not be distinct.

Let a given generic point O on V_3 be taken as origin and the tangent S_3 to V_3 at this point be taken as the coördinate space $w=0$. Let the equation of V_3 be given in each of the forms

(1) $$w=\phi_2(x,y,z)+\phi_3(x,y,z)+\phi_4(x,y,z)+\cdots$$

and

(2) $$w=A(x,y)+zB(x,y)+z^2C(x,y)+z^3D(x,y)+\cdots,$$

where ϕ_i is homogeneous of degree i in x, y, z and A, B, C, etc. are power series in x and y only.

We shall suppose the coördinate system chosen so that no generator in the neighborhood of the origin lies in a space $z=\text{const}$. Let the six systems of rectilinear generators on V_3 be defined on the variety by the equations

(3) $$x=u+z\sigma_i(u,v) \quad y=v+zt_i(u,v) \qquad (i=1,2,3,4,5,6)$$

where

(4) $$\begin{aligned}\sigma_i(u_1v)&=\sigma_0^{(i)}+\sigma_1^{(i)}+\sigma_2^{(i)}+\sigma_3^{(i)}+\cdots\\ \tau_1(u_1v)&=t_0^{(i)}+t_1^{(i)}+t_2^{(i)}+t_3^{(i)}+\cdots \qquad (i=1,2,3,4,5,6)\end{aligned}$$

* Compare, e.g., the author, "On Three-spread Satisfying Four or More Homogeneous, Linear, Partial Differential Equations of the Second Order," *American Journal of Mathematics*, Vol. 33 (1911), pp. 97-128.

in which $\sigma_0^{(i)}$, $t_0^{(i)}$ are constants and $\sigma_j^{(i)}$, $t_j^{(i)}$ are homogeneous of degree j in u and v. From (2) and (3) we have, further, for the generators of each system

$$(5) \qquad w = A(u, v) + z[B(u, v) + A_u\sigma_i + A_v t_i]$$

and the restrictive conditions on the coefficients in (2)

$$(6a) \quad C + B_u\sigma_i + B_v t_i + (1/2!)[A_{uu}\sigma_i^2 + 2A_{uv}\sigma_i t_i + A_{vv}t_i^2] \equiv 0$$
$$(6b) \quad D + C_u\sigma_i + C_v t_i + (1/2!)[B_{uu}\sigma_i^2 + \cdots] + (1/3!)[A_{uuu}\sigma_i^3 + \cdots] \equiv 0$$
$$(6c) \quad E + D_u\sigma_i + D_v t_i + (1/2!)[C_{uu}\sigma_i^2 + \cdots] + \cdots \equiv 0$$
$$(6d) \quad F + E_u\sigma_i + E_v t_i + (1/2!)[D_{uu}\sigma_i^2 + \cdots] + \cdots \equiv 0$$

etc.

Each rectilinear generator at the origin lies in $w = 0$ and also satisfies the equations

$$\phi_2 = 0 \qquad \phi_3 = 0 \qquad \phi_4 = 0 \qquad \text{etc.}$$

In particular, since the six distinct intersections of $\phi_2 = 0$ $\phi_3 = 0$ lie on $\phi_4 = 0$, we have

$$(7) \qquad \phi_4 = \alpha_1\phi_3 + \beta_2\phi_2,$$

where α_1 is linear and β_2 is quadratic in x, y, z.

Consider, now, the cubic hypersurface

$$(8) \quad w = \phi_2 + \phi_3 + \alpha_1(w - \phi_2) + \beta_2 w + \gamma_0 w(w - \phi_2) + \beta_1 w^2 + \beta_0 w^3 = 0$$

where α_1 and β_2 are the forms defined in (7), the quantities β_0, γ_0 are constants, as yet undefined, and β_1 is an undetermined linear form in x, y, z.

We shall show that the quintic cone

$$(9) \qquad \phi_5 - \alpha_1\phi_4 - \beta_2\phi_3 = 0$$

in $w = 0$ has each of the generators $w = 0$ $\phi_2 = 0$ $\phi_3 = 0$ of V_3 at the origin as a double generator. It will then follow that this quintic has $\phi_2 = 0$ as a component and that the residual cubic component passes through each of these generators and is thus of the form $\gamma_0\phi_3 + \beta_1\phi_2 = 0$, so that

$$(10) \qquad \phi_5 \equiv \alpha_1\phi_4 + \beta_2\phi_3 + \gamma_0\phi_2\phi_3 + \beta_1\phi_2^2.$$

If, in (8), we take for γ_0 and β_1 the quantities defined by (10), the expansion in series of the cubic hypersurface (8) will coincide with (1) to the terms of ϕ_5 inclusive.

To show that the cone (9) has the given lines as generators, we take the coördinate system so that any one of the generators at the origin is $w = x = y = 0$ and that

$$\phi_2 \equiv xz - y^2,$$
$$\phi_3 \equiv yz^2 + zB_2(x, y) + A_3(x, y),$$

where B_2 and A_3 denote respectively the second degree terms of B and the third degree terms of A in (2). Let, also

$$\alpha_1(xyz) \equiv \alpha_{00}z + \alpha_{10}x + \alpha_{01}y,$$
$$\beta_2(x_1y_1z) \equiv \beta_{00}z^2 + \beta_{10}xz + \beta_{01}yz + \cdots;$$

then, from (7),

$$\phi_4 \equiv \beta_{00}xz^3 + \alpha_{00}yz^3 + \cdots.$$

Since $x = y = w = 0$ is a generator, ϕ_5 vanishes for $x = y = 0$. Let

$$\phi_5 \equiv a_5xz^4 + b_5yz^4 + \cdots.$$

From (6a) and (6b), since $\sigma_{00} = t_{00} = 0$, we have

(11) $$\sigma_1 \equiv -v, \qquad t_1 \equiv -(\beta_{00}u + \alpha_{00}v),$$

where σ_1 and t_1 are the first degree terms of σ and t respectively for the system that contains the line $x = y = w = 0$.

From (6c) we now have

$$\phi_5(u, v, 1) \equiv \alpha_{00}\beta_{00}u + (\beta_{00} + \alpha_{00}^2)v + \cdots.$$

But, in coördinates $u, v, 1$, we have

$$\alpha_1\phi_4 + \beta_2\phi_3 \equiv \alpha_{00}\beta_{00}u + (\beta_{00} + \alpha_{00}^2)v + \cdots.$$

Hence the cone $\phi_5 - (\alpha_1\phi_4 + \beta_2\phi_3) = 0$ in $w = 0$ has the generator $x = y = w = 0$ as a double line, and, since this is any one of the six generators at the origin, ϕ_5 is of the form (10).

Next, we shall show that the cone

(12) $$\phi_6 - [\alpha_1\phi_5 + \beta_2\phi_4 + \gamma_0\phi_2\phi_4 + \gamma_0\phi_3^2 + 2\beta_1\phi_2\phi_3] = 0$$

in $w = 0$ has each of the generators $w = \phi_2 = \phi_3 = 0$ of V_3 at the origin as a triple line from which it will follow at once that

(13) $$\phi_6 \equiv \alpha_1\phi_5 + \beta_2\phi_4 + \gamma_0\phi_2\phi_4 + \gamma_0\phi_3^2 + 2\beta_1\phi_2\phi_3 + \beta_0\phi_2^3.$$

If, in (9), we take for β_0 the value determined by (13), then the expansion of the cubic hypersurface coincides with that of V_3 as given by (1) to the terms of ϕ_6 inclusive.

If we denote by A_2, B_2, C_2, etc., the terms of A, B, C, etc., of second degree in u and v, we find, from (6a) and (6b)

$$\sigma_2 \equiv -[C_2 + B_{2u}\sigma_1 + B_{2v}t_1 + A_2(\sigma_1t_1)]$$
$$t_2 \equiv -[D_2 + C_{2u}\sigma_1 + C_{2v}t_1 + B_2(\sigma_1t_1)]$$

from which, using the values of σ_1, t_1 determined by (11), we have σ_2, t_2

determined uniquely in terms of the coefficients of ϕ_2, ϕ_3, ϕ_4, ϕ_5. From (6c), (6d), and (6e), we now have

$$E_2 \equiv -[D_{1u}\sigma_2 + D_{1v}t_2 + D_{2u}\sigma_1 + D_{2v}t_1 + C_2(\sigma_1 t_1)]$$
$$F_1 \equiv [E_{1u}\sigma_1 + E_{1v}t_1], \qquad G_0 = 0.$$

These equations fix the coefficients of the terms of $\phi_6(u, v, 1)$ of degree not more than two in u and v uniquely in terms of those of ϕ_2, ϕ_3, ϕ_4, ϕ_5. It follows that these coefficients are identical with those of the corresponding expansion of the cubic (8). But, for (8), the sextic corresponding to (12) is identical with $-\beta_0\phi_2{}^3 = 0$ and thus has a triple generator along $x = y = w = 0$. It follows that the given sextic (12) has each generator of $w = \phi_2 = \phi_3 = 0$ as a triple generator and thus is identical with a constant times $\phi_2{}^3 = 0$. Hence ϕ_6 is of the form given by equation (13).

Finally, the terms of $\phi_7(u, v, 1)$ involving u and v to not more than the third power are determined uniquely by those of $\phi_2, \cdots, \phi_6$. In fact, from (6a), (6b), we obtain

$$\sigma_3 \equiv -[C_3 + B_{2u}\sigma_2 + B_{3u}\sigma_1 + B_{2v}t_2 + B_{3v}t_1 - 2t_1t_2 + \tfrac{1}{2}(A_{3uu}\sigma_1{}^2 + 2A_{3uv}\sigma_1 t_1 + A_{3vv}t_1{}^2)]$$
$$t_3 \equiv -[D_3 + C_{2u}\sigma_2 + C_{3u}\sigma_1 + C_{2v}t_2 + C_{3v}t_1 + B_{2uu}\sigma_1\sigma_2 + B_{2uv}(\sigma_1 t_2 + \sigma_2 t_1) + B_{2vv}t_1t_2 + \tfrac{1}{2}(B_{3uu}\sigma_1{}^2 + 2B_{3uv}\sigma_1 t_1 + B_{3vv}t_1{}^2) + A_3(\sigma_1, t_1)].$$

From these equations, σ_3 and t_3 are uniquely determined in terms of the coefficients of $\phi_2, \cdots, \phi_6$.

From (6c), (6d), (6e) and (6f), we find

$$E_3 \equiv -[D_{1u}\sigma_3 + D_{2u}\sigma_2 + D_{3u}\sigma_1 + D_{1v}t_3 + D_{2v}t_2 + D_{3v}t_1 + C_{2uu}\sigma_1\sigma_2 + C_{2uv}(\sigma_1 t_2 + \sigma_2 t_1) + C_{2vv}t_1t_2 + \tfrac{1}{2}(C_{3uu}\sigma_1{}^2 + 2C_{3uv}\sigma_1 t_1 + C_{3vv}t_1{}^2) + A_3(\sigma_1 t_1)]$$
$$F_2 \equiv -[E_{1u}\sigma_2 + E_{2u}\sigma_1 + E_{1v}t_2 + E_{2v}t_1 + D_2(\sigma_1 t_1)]$$
$$G_1 \equiv -[F_{1u}\sigma_1 + F_{1v}t_1] \qquad H_0 = 0.$$

These equations fix the terms of $\phi_7(u, v, 1)$ of degree three or less in u and v terms of those of $\phi_2, \cdots, \phi_6$. It follows that these are identical with the corresponding terms of ϕ_7', which arises from the seventh degree terms of the cubic (8) so that $\phi_7(x, y, z) - \phi_7'(x, y, z) = 0$ has $x = y = w = 0$ as a fourfold line. Since this holds for each generator $w = \phi_2 = \phi_3 = 0$, we have $\phi_7 \equiv \phi_7'$. Thus the cubic (8) has contact with (1) to the terms in ϕ_7 inclusive. Since this contact holds at an origin chosen generically on V_3, it follows that the given variety coincides with the cubic hypersurface.

The Oscillation of a Sequence.*

By WALLIE ABRAHAM HURWITZ.

1. *Introduction.* For any given sequence (x) of real or complex numbers $x_1, x_2, x_3, \cdots$ let

$$\Omega(x) = \limsup_{m\to\infty,\ n\to\infty} |x_m - x_n|$$

be termed the *oscillation* of the sequence. In order that (x) be convergent it is necessary and sufficient that $\Omega(x) = 0$. The oscillation of a sequence may be considered as measuring the deviation of the sequence from convergence, just as the oscillation of a function at a point may be considered as measuring the deviation of the function from approach to a limit at the point. As defined, $\Omega(x)$ may be a (finite) number $\geqq 0$ or the symbol $+\infty$.

In order that $\Omega(x)$ be finite, it is necessary and sufficient that (x) be bounded.

If (x) is bounded, obviously $\Omega(x)$ is finite. On the other hand suppose that $\Omega(x)$ is finite, and let $Q > \Omega(x)$. There exists an integer q such that when $m \geqq q$, $n \geqq q$, then $|x_m - x_n| < Q$. In particular, for $n \geqq q$, $|x_n - x_q| < Q$ and $|x_n| < Q + |x_q|$; therefore the greatest of the q numbers

$$|x_1|,\ |x_2|, \cdots, |x_{q-1}|,\ Q + |x_q|$$

is an upper bound of all $|x_n|$.

For a bounded sequence, $\Omega(x)$ may also be defined as the maximum of all distances between pairs of limiting values of (x). For a bounded real sequence

(1) $$\Omega(x) = \limsup_{n\to\infty} x_n - \liminf_{n\to\infty} x_n.$$

2. *Effect of regular transformation.* By the familiar Silverman-Toeplitz theorem, in order that the transformation

(A) $$y_n = \sum_{k=1}^{n} a_{n,k} x_k$$

be regular, i. e., transform every convergent sequence (x) into a sequence (y) having the same limit, it is necessary and sufficient that

* Read before the American Mathematical Society, December 27, 1929.

$$(2)\qquad \begin{cases} \text{(I)} & \textit{for each } k, \lim_{n\to\infty} a_{n,k} = 0; \\ \text{(II)} & \lim_{n\to\infty} \sum_{k=1}^{n} a_{n,k} = 1; \\ \text{(III)} & \textit{for all } n, \sum_{k=1}^{n} |a_{n,k}| \textit{ is bounded.} \end{cases}$$

The object of the present paper is to consider the change eff oscillation of a sequence by a regular linear transformation of the and specifically to insure that

$$(3)\qquad \Omega(y) \leqq \Omega(x).$$

The problem has an obvious relation to a result obtained in an ear it was shown that if the elements of (A) are real, and only real se are considered, then in order that a regular transformation (such that

$$(4)\qquad \lim_{n\to\infty} \sup y_n \leqq \lim_{n\to\infty} \sup x_n$$

and

$$(5)\qquad \lim_{n\to\infty} \inf y_n \geqq \lim_{n\to\infty} \inf x_n$$

whenever (x) is bounded, it is necessary and sufficient that

$$(6)\qquad \text{(IV)} \quad \lim_{n\to\infty} \sum_{k=1}^{n} |a_{n,k}| = 1.$$

In fact, for a bounded real sequence, it is clear that (IV), bei for (4) and (5), is sufficient for (3). But (3) does not imply l (5), so that the theorem mentioned does not show that (IV) for (3). A slight modification of the proof given would es necessity, even for the complex case; but in the complex case, c hand, (4) and (5) are meaningless and therefore cannot be us to show sufficiency. The whole question is therefore now dis pendently; the following theorem will be proved:

In order that a regular transformation (A) *may be such that* Ω *for every sequence* (x), (IV) *is necessary and sufficient.*

3. *Proof of Necessity.* To prove the necessity of (IV) we s (I), (II), (III) hold, but (IV) fails, and exhibit a sequence $(x$ $\Omega(y) > \Omega(x)$, thereby establishing a contradiction.† If

* W. A. Hurwitz, *Proceedings of the London Mathematical Socie* Vol. 26 (1927), pp. 231-248.

† The construction of the sequence (x) involves only minor modific proof previously mentioned; *loc. cit.*, pp. 238-239.

$$\limsup_{n\to\infty} \sum_{k=1}^{n} | a_{n,k} | > 1,$$

then for some $\alpha > 1$, there will exist values of n arbitrarily great for which $\sum_{k=1}^{n} | a_{n,k} | > \alpha$. Choose n_1 so that

$$\sum_{k=1}^{n_1} | a_{n_1,k} | > \alpha.$$

By (1), $\sum_{k=1}^{n_1} | a_{n,k} | \to 0$ as $n \to \infty$; we can therefore choose n_2 so that

$$\sum_{k=1}^{n_1} | a_{n_2,k} | < 1/2, \qquad \sum_{k=1}^{n_2} | a_{n_2,k} | > \alpha.$$

In general, choose n_p so that

$$\sum_{k=1}^{n_{p-1}} | a_{n_p,k} | < 1/p, \qquad \sum_{k=1}^{n_p} | a_{n_p,k} | > \alpha;$$

then also

$$\sum_{k=n_{p-1}+1}^{n_p} | a_{n_p,k} | > \alpha - 1/p.$$

Now define *

$$x_k = \begin{cases} \operatorname{sgn} a_{n_1,k}, & k \leqq n_1; \\ (-1)^{p-1} \operatorname{sgn} a_{n_p,k}, & n_{p-1} < k \leqq n_p. \end{cases}$$

Then for all k, $| x_k | \leqq 1$, and $\Omega(x) \leqq 2$. We find that

$$\begin{aligned} y_{n_{p+1}} - y_{n_p} &= \sum_{k=1}^{n_{p+1}} a_{n_{p+1},k} x_k - \sum_{k=1}^{n_p} a_{n_p,k} x_k \\ &= \sum_{k=1}^{n_p} a_{n_{p+1},k} x_k - \sum_{k=1}^{n_{p-1}} a_{n_p,k} x_k + \sum_{k=n_p+1}^{n_{p+1}} a_{n_{p+1},k} x_k - \sum_{k=n_{p-1}+1}^{n_p} a_{n_p,k} x_k \\ &= \sum_{k=1}^{n_p} a_{n_{p+1},k} x_k - \sum_{k=1}^{n_{p-1}} a_{n_p,k} x_k \\ &\qquad + (-1)^p \sum_{k=n_p+1}^{n_{p+1}} | a_{n_{p+1},k} | + (-1)^p \sum_{k=n_{p-1}+1}^{n_p} | a_{n_p,k} | . \end{aligned}$$

In the last expression the absolute values of the first two terms are together not greater than

$$\sum_{k=1}^{n_p} | a_{n_{p+1},k} | + \sum_{k=1}^{n_{p-1}} | a_{n_p,k} | < 1/(p+1) + 1/p < 2/p;$$

the absolute values of the last two terms (which are real and of like sign) are together greater than

* Here sgn $z = |z|/z$ or 0 according as $z \neq 0$ or $z = 0$; hence in all cases z sgn $z = |z|$.

$$\{\alpha-[1/(p+1)]\}+[\alpha-(1/p)]>2\alpha-2/p.$$

Hence

$$|y_{n_{p+1}}-y_{n_p}|>2\alpha-4/p.$$

Thus

$$\Omega(y)=\limsup_{m\to\infty,\,n\to\infty}|y_m-y_n|\geqq\limsup_{p\to\infty}|y_{n_{p+1}}-y_{n_p}|$$
$$\geqq\lim_{p\to\infty}(2\alpha-4/p)=2\alpha>2\geqq\Omega(x),$$

and $\Omega(y)>\Omega(x)$.

4. *Proof of Sufficiency.* It will now be shown that if (I), (II), (III), (IV) hold, then $\Omega(y)\leqq\Omega(x)$. In case $\Omega(x)=+\infty$, no proof is necessary. Let $\Omega(x)=\omega$ be finite; since (x) must be bounded, let $|x_n|\leqq X$ for all n; let $\sum_{k=1}^{n}|a_{n,k}|=A_n$. For $\epsilon>0$ there is an integer p such that when $\mu>p$, $\nu>p$, then $|x_\mu-x_\nu|<\omega+\epsilon$. If $m>p$, $n>p$,

$$y_m-y_n=\sum_{k=1}^{m}a_{m,k}x_k-\sum_{k=1}^{n}a_{n,k}x_k$$
$$=\sum_{k=1}^{p}(a_{m,k}-a_{n,k})x_k+\sum_{k=p+1}^{m}a_{m,k}x_k-\sum_{k=p+1}^{n}a_{n,k}x_k.$$

The absolute value of the first term of the last expression is not greater than

$$2X(\sum_{k=1}^{p}|a_{m,k}|+\sum_{k=1}^{p}|a_{n,k}|,$$

which, by (I), has the limit 0 as m and n become infinite. Thus

$$\Omega(y)=\limsup_{m\to\infty,\,n\to\infty}|\sum_{k=p+1}^{m}a_{m,k}x_k-\sum_{k=p+1}^{n}a_{n,k}x_k|.$$

Furthermore

$$\sum_{k=p+1}^{m}a_{m,k}x_k-\sum_{k=p+1}^{n}a_{n,k}x_k=\sum_{\mu=p+1}^{m}a_{m,\mu}x_\mu-\sum_{\nu=p+1}^{n}a_{n,\nu}x_\nu$$
$$=(\sum_{\mu=p+1}^{m}a_{m,\mu}x_\mu)(1-\sum_{\nu=p+1}^{n}a_{n,\nu})-(\sum_{\nu=p+1}^{n}a_{n,\nu}x_\nu)(1-\sum_{\mu=p+1}^{m}a_{m,\mu})$$
$$+\sum_{\mu=p+1}^{m}\sum_{\nu=p+1}^{n}a_{m,\mu}a_{n,\nu}(x_\mu-x_\nu).$$

But

$$|(\sum_{\mu=p+1}^{m}a_{m,\mu}x_\mu)(1-\sum_{\nu=p+1}^{n}a_{n,\nu})|\leqq XA_m|1-\sum_{\nu=p+1}^{n}a_{n,\nu}|,$$

$$|(\sum_{\nu=p+1}^{n}a_{n,\nu}x_\nu)(1-\sum_{\mu=p+1}^{m}a_{m,\mu})|\leqq XA_n|1-\sum_{\mu=p+1}^{n}a_{m,\mu}|,$$

each of which, by (I), (II), and (IV), has the limit 0 as m and n become infinite. Thus

$$\Omega(y) = \limsup_{m\to\infty,\, n\to\infty} \left| \sum_{\mu=p+1}^{m} \sum_{\nu=p+1}^{n} a_{m,\mu} a_{n,\nu} (x_\mu - x_\nu) \right| .$$

Finally,

$$\left| \sum_{\mu=p+1}^{m} \sum_{\nu=p+1}^{n} a_{m,\mu} a_{n,\nu} (x_\mu - x_\nu) \right| \leqq \sum_{\mu=p+1}^{m} \sum_{\nu=p+1}^{n} | a_{m,\mu} || a_{n,\nu} | (\omega + \epsilon)$$
$$\leqq (\omega + \epsilon) A_m A_n,$$

so that, by (IV)

$$\Omega(y) \leqq \omega + \epsilon.$$

Since this holds for each $\epsilon > 0$, it follows that $\Omega(y) \leqq \omega = \Omega(x)$.

5. *Variations of the Theorem.* For the special sequence used in § 3 in order to prove the necessity of (IV), $\Omega(x)$ is finite; hence

In order that the regular transformation (A) *may be such that* $\Omega(y) \leqq \Omega(x)$ *for every sequence* (x) *such that* $\Omega(x)$ *is finite,* (IV) *is necessary and sufficient.*

The Silverman-Toeplitz theorem holds if we restrict ourselves to real transformations of real sequences; also if every element of (A) is real, then the sequence (x) chosen in § 3 is real; hence

In order that the regular transformation (A) *with real elements may be such that* $\Omega(y) \leqq \Omega(x)$ *for every real sequence* (x), (IV) *is necessary and sufficient.*

6. *An Interpretation.* It may not be superfluous to point out that when (II) is satisfied, (IV) is equivalent to the requirement that in the sum of absolute values of elements of a row of (A), the total contribution of the pure imaginary parts and of those real parts which are negative, is negligible. In order to state this more precisely and prove it, we note the following lemma:

For any real numbers u_k, v_k,

$$\left(\sum_{k=1}^{n} u_k\right)^2 + \left(\sum_{k=1}^{n} v_k\right)^2 \leqq \left[\sum_{k=1}^{n} (u_k^2 + v_k^2)^{1/2}\right]^2.$$

For $n = 2$, this follows at once by rationalization; for $n > 2$, it can be shown by induction or repeated application of the case $n = 2$.

In this formula write $u_k = | b_{n,k} |$, $v_k = | c_{n,k} |$, where $a_{n,k} = b_{n,k} + i c_{n,k}$ and $b_{n,k}$, $c_{n,k}$ are real:

$$(7) \qquad \left(\sum_{k=1}^{n} | b_{n,k} | \right)^2 + \left(\sum_{k=1}^{n} | c_{n,k} | \right)^2 \leqq \left(\sum_{k=1}^{n} | a_{n,k} | \right)^2.$$

Suppose (II) holds; then

$$\lim_{n\to\infty} \sum_{k=1}^{n} b_{n,k} = 1 \tag{8}$$

and

$$\sum_{k=1}^{n} b_{n,k} \leqq \sum_{k=1}^{n} | b_{n,k} | \leqq \sum_{k=1}^{n} | a_{n,k} | . \tag{9}$$

If (IV) also holds, then by (IV), (8), and (9),

$$\lim_{n\to\infty} \sum_{k=1}^{n} | b_{n,k} | = 1; \tag{10}$$

and by (IV), (7), and (10),

$$\lim_{n\to\infty} \sum_{k=1}^{n} | c_{n,k} | = 0. \tag{11}$$

But (IV) can also be deduced from (II), (10), and (11), since

$$\sum_{k=1}^{n} a_{n,k} \leqq \sum_{k=1}^{n} | a_{n,k} | \leqq \sum_{k=1}^{n} | b_{n,k} | + \sum_{k=1}^{n} | c_{n,k} | .$$

Thus:

If (II) *holds,* (IV) *is equivalent to* (10) *and* (11). Now let

$$b^{+}_{n,k} = (b_{n,k} + | b_{n,k} |)/2, \quad b^{-}_{n,k} = (b_{n,k} - | b_{n,k} |)/2, \tag{12}$$

so that

$$b_{n,k} = \begin{cases} b^{+}_{n,k} & \text{if} \quad b_{n,k} > 0, \\ b^{-}_{n,k} & \text{if} \quad b_{n,k} < 0; \end{cases}$$
$$b^{+}_{n,k} \geqq 0; \qquad b^{-}_{n,k} \leqq 0.$$

Then

$$b_{n,k} = b^{+}_{n,k} + b^{-}_{n,k}; \qquad | b_{n,k} | = b^{+}_{n,k} - b^{-}_{n,k}. \tag{13}$$

From (10) and (8), by (12), we find

$$\lim_{n\to\infty} \sum_{k=1}^{n} b^{+}_{n,k} = 1, \qquad \lim_{n\to\infty} \sum_{k=1}^{n} b^{-}_{n,k} = 0; \tag{14}$$

while from (14), by (13), we could deduce (10) and (8); from (8) and either equation of (14), by (13), we could deduce (10). Thus:

(II) *and* (IV) *are equivalent to* (11) *and* (14). *If* (II) *holds, then* (IV) *is equivalent to* (11) *and either equation of* (14).

Determination of All the Groups of Order 64.

By G. A. Miller.

Introduction.

Among all the groups whose orders do not exceed 100 those of order 64 present by far the greatest difficulties in view of the very large number of distinct groups of this order. The groups of every other order within this limit have been published either separately or as special cases of well known general categories. In fact, since $64 = 2^6$ these groups are also included among those of order p^6, p being a prime number. This general category is the subject of a Paris thesis, 1904, by M. Potron, but not all the possible groups of order 64 appear therein, and the more general methods of treatment which are employed there naturally involve many difficulties which do not present themselves in this special category. It may be of interest to tabulate here the number of groups of every order not greater than 100 whenever there is more than one group of the same order.

Order	4	6	8	9	10	12	14	16	18	20	21	22	24	25	26	27
Number	2	2	5	2	2	5	2	14	5	5	2	2	15	2	2	5
Order	28	30	32	34	36	38	39	40	42	44	45	46	48	49	50	52
Number	4	4	51	2	14	2	2	14	6	4	2	2	52	2	5	5
Order	54	55	56	57	58	60	62	63	64	66	68	70	72	74	75	76
Number	15	2	13	2	2	13	2	4	294	4	5	4	50	2	3	4
Order	78	80	81	82	84	86	88	90	92	93	94	96	98	99	100	
Number	6	52	15	2	15	2	12	10	4	2	2	230	5	2	16	

The total number of groups whose orders do not exceed 100 is therefore 1074 and for each of 37 of these orders there is only one group.

1. *Abelian groups.* An abelian group of order p^m, p being a prime number, whose invariants are $p^{m_1}, p^{m_2}, \cdots, p^{m_\lambda}$ is said to be of type $(m_1, m_2, \cdots, m_\lambda)$, and there is only one such group if $m_1, m_2, \cdots, m_\lambda$ represent positive numbers such that $m_1 + m_2 + \cdots + m_\lambda = m$, and such a group is completely defined by its type. Hence there are exactly eleven groups of order 2^6, which correspond respectively to the following partitions of 6 with respect to addition: 6, $5+1$, $4+2$, $4+1+1$, $3+3$, $3+2+1$, $3+1+1+1$, $2+2+2$, $2+2+1+1$, $2+1+1+1+1$, $1+1+1+1+1+1$. It is known that every non-abelian group of order 64 contains an invariant abelian subgroup of order 16, but it does not necessarily contain an abelian subgroup of order 32.[*] We may therefore divide the non-abelian

* Miller, Blichfeldt, Dickson; *Finite Groups* (1916), p. 120.

groups of order 64 into two categories which are composed respectively of all those groups which involve at least one abelian subgroup of index 2 and of all those which do not have this property. The former category will be considered first.

2. *General theorems relating to non-abelian groups which contain an abelian subgroup of index* 2. If H represents an abelian subgroup of index 2 under the group G, then G can be obtained by adjoining to H an operator s which transforms H according to an automorphism of order 2 and has its square in H, and any operator which satisfies these two conditions gives rise to a group whose order is twice the order of H when it is adjoined to H. In order to determine these automorphisms of order 2 it is desirable to note that every commutator which arises from an automorphism of order 2 of any group whatever corresponds to its inverse under this automorphism, and that the commutators which arise from an automorphism of an abelian group constitute a subgroup of this abelian group.

Since every commutator which results from an automorphism of order 2 gives rise to a commutator which is the inverse of the square of the former commutator, it results that, when the order of a commutator arising from an automorphism of order 2 of an abelian group is odd, then every generator of this commutator which appears in this abelian group and has an order involving no prime factors besides those of the order of the commutator must give rise under the same automorphism to a commutator whose order is equal to the order of this generator. As this commutator must correspond to its inverse under this automorphism, there results the following theorem: *If a commutator of odd order results from an automorphism of order 2 of an abelian group, and if s is any generator of this commutator whose order involves no prime factor besides those which divide the order of this commutator, then an operator whose order is equal to that of s corresponds to its inverse under the same automorphism.* From this theorem it results directly that in an automorphism of order 2 of a cyclic group of odd prime power order every operator must correspond to its inverse, and hence there is only one possible automorphism of order 2 of such a cyclic group.

When a commutator arising from an automorphism of order 2 of an abelian group is of even order, this commutator will give rise to a commutator whose order is one-half of the order of the former commutator. In particular, when a commutator of order 2 arises from such an automorphism, it must correspond to itself under this automorphism. In general, a generator of a commutator of even order arising from such an automorphism must give rise to a commutator whose order is equal to one-half the order of this generator

whenever the order of this generator involves no prime factors except those which appear in the order of the commutator. From this it results that when the abelian group is cyclic and of order 2^m, then there is one automorphism of order 2 which gives rise to a commutator of order 2, while in the other possible automorphisms of order 2 at least one-half of the operators of this cyclic group must correspond to their inverses. There can therefore be only two such other automorphisms. Since every abelian group admits an automorphism in which every operator corresponds to its inverse, and such an automorphism of a cyclic group of order 2^m can not be the square of another automorphism of this group, it results from the given theorems that the group of isomorphisms of this cyclic group is of type $(\alpha, 1)$. This is a new proof of this well known fact and illustrates the usefulness of the theorems noted above.

In view of the fact that the multiplying group can often be used to simplify the study of the automorphisms of an abelian group, we note here a few more properties of this group. It is especially useful to know some of the properties of the multiplying group which corresponds to the transform of a given automorphism under the group of isomorphisms. It may first be noted that if under an automorphism of an abelian group the multiplier of a given operator s is t, then under the inverse of this automorphism t^{-1} is the multiplier of ts. In particular, *the multiplying group corresponding to an automorphism is the same as that which corresponds to the inverse of this automorphism.* A necessary and sufficient condition that the multiplier of an operator s under a given automorphism is the inverse of the multiplier of s under the inverse of this automorphism is that this multiplier is invariant under this automorphism. In general, if the multipliers of s corresponding to two automorphisms of an abelian group are t_1 and t_2, then the multiplier of s in the product of these two automorphisms, in order, is $t_3t_2t_1$, where t_3 is the multiplier of t_1 under the second automorphism.

The preceding theorems may be illustrated by a consideration of all the possible automorphisms of order 2 of the abelian group H of order 2^α and of type $(\alpha - 1, 1)$, $\alpha > 4$. The commutator subgroup corresponding to such an automorphism must be cyclic since a non-cyclic commutator subgroup would involve all the operators of order 2 contained in H and hence all these operators would correspond to themselves in this automorphism. The quotient group corresponding to a subgroup involving all the operators of order 2 contained in H is obviously cyclic. There are three non-conjugate operators of order 2 in the group of isomorphisms of H which correspond to commutator subgroups of order 2. When the commutator subgroup is of order 4, it is

not contained in a cyclic subgroup of order 8, and hence the two corresponding automorphisms of order 2 are conjugate.

When the order of the commutator subgroups corresponding to an automorphism of order 2 exceeds 4 it must be $2^{\alpha-2}$ since at least one-half of the operators of H must correspond to their inverses in such an automorphism in view of the fact that an operator which is the square of an operator of highest order must correspond to its inverse therein. There are two cyclic subgroups of order $2^{\alpha-2}$ in H and each of these is a commutator subgroup arising from an automorphism of order 2. When the commutator subgroup of this order is generated by the square of an operator of highest order contained in G and the corresponding central is non-cyclic, every operator of H corresponds either to its inverse or to its $2^{\alpha-2}-1$ power. These two automorphisms are not conjugate under the group of isomorphisms of H, but the two automorphisms of order 2 which arise when the central is cyclic and the commutator subgroup remains unchanged are conjugate. When the commutator subgroup is the other cyclic subgroup of order $2^{\alpha-2}$, the four possible automorphisms are conjugate in pairs. Hence the following theorem: *The group of isomorphisms of the abelian group of order 2^α and of type $(\alpha-1, 1)$, $\alpha > 4$, involves fifteen operators of order 2; three of these are invariant and the other twelve are conjugate in pairs. Hence there are nine such automorphisms which have the property that no two of them are conjugate under the group of isomorphisms.*

A very simple illustration of the general theorems noted above is furnished by the operations of order 2 in the group of isomorphisms of the abelian group of order 2^α and of type $(1, 1, 1 \cdots)$. In this case all the commutators must correspond to themselves in such an automorphism, and hence the number of such non-conjugate automorphisms is equal to the largest integer which does not exceed $\alpha/2$. When the abelian group is of order 2^α and of type $(2, 1, 1 \cdots)$, the commutator subgroup which corresponds to an automorphism of order 2 must again appear in the central under the automorphism, and the number of the independent generators of this commutator subgroup may be any natural number which does not exceed $\alpha/2$. When the commutator subgroup is of order 2 and $\alpha > 3$, there are always four such automorphisms which are not conjugate under the group of isomorphisms. The largest number of non-conjugate automorphisms of order 2 corresponding to a given order of the commutator subgroup is obviously 5, and this number is always attained when this order exceeds two and its square is less than one-half of the order of the group. When this square is equal to one-half of the order of the group, there are four such automorphisms, and when it is

equal to the order of the group, there are two such automorphisms. This gives rise to the following theorem: *The number of non-conjugate operators of order 2 in the group of isomorphisms of the abelian group of order* 2^m *and of type* $(2, 1, 1 \cdots)$ *is one if* $m = 2$, *three if* $m = 3$, *and if* $m > 3$ *and is even it is* $1 + 5(m - 2)/2$, *while when m is odd it is* $3 + 5(m - 3)/2$.

3. *Groups which involve either the cyclic group of order* 32 *or the abelian group of order* 32 *and of type* $(4, 1)$. It is well known that there are always four non-abelian groups of order 2^m, $m > 3$, which involve the cyclic group of order 2^{m-1}. As these four groups are so well known, we shall not consider their properties here. It was noted in the preceding section that when H is the abelian group of order 32 and of type $(4, 1)$ there are nine non-conjugate automorphisms of order 2, and hence we proceed to consider the possible groups of order 64 which may correspond to these automorphisms.

When the commutator subgroup is the characteristic subgroup of order 2 and the central is cyclic, there is only one group which does not involve an operator of order 32, while there are two groups when this central is non-cyclic. When the commutator subgroup is a non-characteristic subgroup of order 2, the central must be non-cyclic and there are two additional groups. Hence five distinct groups of order 64 result when the commutator subgroup is of order 2. Each of these five groups contains three abelian subgroups of index 2 in view of the obvious theorem that a necessary and sufficient condition that a non-abelian group of order p^m which involves an abelian subgroup of index p contains $p + 1$ abelian subgroups of this index is that the commutator subgroup is of order p.

When the commutator subgroup of H is of order 4, the central must be cyclic and there is only one corresponding group of order 64. When the commutator subgroup is of order 8, the central may be either cyclic or non-cyclic and the commutator subgroup may be either of the two cyclic subgroups of order 8. Hence there are four such automorphisms. There are three groups when each operator of H corresponds to its inverse, and there are two groups when these operators correspond to their seventh powers. When the central is non-cyclic and the commutator subgroup is the other cyclic subgroup of order 8, there are two additional groups. Hence there are seven groups when the central is non-cyclic and the commutator subgroup is of order 8. When the central is cyclic, it is the subgroup of order 4 which cannot be generated by an operator of order 8 contained in H. There are two groups when the commutator subgroup is the cyclic group which is generated by an operator of order 16, while there is only one group when the commutator subgroup is the other group of order 8. Hence there are 16

groups of order 64 which contain the abelian subgroup of type (4, 1) but no operator of order 32.

4. *Groups which involve the abelian group of type* (3, 2) *but no abelian group of order* 32 *which involves an operator of order* 16. In view of the theorem that if a non-abelian group of order p^m contains more than one abelian subgroup of index p, then each of its operators appears in at least one of its $p+1$ abelian subgroups of this index, it is clear that the restriction imposed on the groups under consideration is equivalent to saying that when the commutator subgroup is of order 2, these groups cannot contain an operator of order 16. There are four non-conjugate automorphisms of order 2 in which the corresponding commutator subgroup is of order 2. When the central involves operators of order 8 and the commutator subgroup is the characteristic subgroup of order 2, there is only one group, but there are two groups when the commutator subgroup is not characteristic. When the central is of type (2, 2), two groups correspond to each of the two non-conjugate automorphisms but only three of these are distinct, and hence there are six groups when H is of type (3, 2) and the commutator subgroup is of order 2. When the commutator subgroup is the non-cyclic group of order 4, the central is of type (2, 1) and there are two non-conjugate automorphisms. Two groups correspond to one of these, while only one group results from the other. Hence three groups of order 64 result from the non-cyclic commutator subgroup of order 4.

When the commutator subgroup is one of the two cyclic groups of order 4 whose operators cannot be used as an independent generator of H, then there are two non-conjugate automorphisms of order 2. In one of these, exactly half of the operators of H correspond to their inverses, and this gives rise to three distinct groups of order 64, while the other gives rise to two additional groups of this order. Only one automorphism of order 2 corresponds to the other cyclic group of order 4 as a commutator subgroup, and this gives rise to two additional groups of order 64. Hence the cyclic commutator subgroups of order 4 give rise to seven groups of order 64 when the central is non-cyclic. It gives rise to two such groups when the central is cyclic. One of these two groups involves operators of order 16 which do not appear in an abelian subgroup of order 32.

It remains to consider the case when the commutator subgroup is of order 8 and hence it is non-cyclic. The central must therefore be the non-cyclic subgroup of order 4 contained in H, and the commutator subgroup must be composed of the operators in H which are squares. There are three groups in which all the operators of H correspond to their inverses, and two

additional ones when all of these operators correspond to their third powers. The remaining two automorphisms in which all the operators of order 4 correspond to their inverses are conjugate and there are two additional groups which correspond to one of these automorphisms. There are only two non-conjugate automorphisms of order 2 when not all the operators of order 4 correspond to their inverses. In one of these one-half of the operators of order 8 correspond to their inverses while the rest correspond to their third powers. Two groups of order 64 arise from this automorphism, while only one group results from the remaining automorphisms. Hence ten distinct groups result when the commutator subgroup is of order 8, and there are 28 groups which result from the automorphisms considered in this section and are not simply isomorphic with the groups resulting from the automorphisms considered in the preceding sections.

5. *Groups which involve the abelian group of type* $(3, 1, 1)$ *but no abelian subgroup of index* 2 *which has less than three independent generators.* Since a group of order 64 which contains the abelian group of type $(3, 1, 1)$ cannot contain any other abelian subgroup of order 32 when the order of its commutator subgroup exceeds 2, it is only necessary to exclude the groups from the enumeration of the present section in which one of the two other abelian subgroups contains just two independent generators. The group of type $(3, 1, 1)$ contains seven subgroups of index 2. One of these is of type $(2, 1, 1)$, and the other six are of type $(3, 1)$ and are conjugate under the group of isomorphisms of H. When the central is of type $(2, 1, 1)$, there are two groups whose commutator subgroup is the characteristic subgroup of order 2, and there are three additional groups whose commutator subgroup is a non-characteristic subgroup of order 2. When the central is of type $(3, 1)$, there is obviously only one additional group. Hence there are 6 groups of order 64 when the commutator subgroup is of order 2.

When the commutator subgroup is the non-cyclic group of order 4, it must appear in the central of H, and hence this central must be of type $(2, 1)$. There are therefore two non-conjugate such automorphisms of order 2. In one of these, four operators of order 4 correspond to their inverses and this gives rise to two groups, while the other gives rise to only one group. Hence three groups of order 64 result when the commutator subgroup is the non-cyclic group of order 4. When the commutator subgroup is the cyclic group of order 4 which is generated by an operator of order 8 and the central is of type $(1, 1, 1)$, every operator of H corresponds either to its inverse or to its third power. In the former case there are three groups while there are two such groups in the latter. When the commutator subgroup remains un-

changed but the central is of type $(2, 1)$, then half of the operators of H correspond to their inverses while the rest correspond to their third powers. There are three such groups. When the commutator subgroup is another cyclic subgroup of order 4 and the central is of type $(1, 1, 1)$, there are three groups and there are two additional groups when the central is of type $(2, 1)$. Hence 16 distinct groups result when the commutator subgroup is of order 4.

It remains to consider the possible groups when the commutator subgroup is of order 8. As this must be non-cyclic, the central must be the non-cyclic group of order 4 which involves the characteristic subgroup of order 2 and an independent generator of this order. The commutator subgroup must involve this central and the subgroup of order 4 which is generated by an operator of order 8. Hence at least 8 operators must correspond to their inverses and there are two non-conjugate automorphisms. In one of these, 16 operators correspond to their inverses, while in the other, eight operators correspond to their third powers. Hence three groups result when the commutator subgroup is of order 8 and the total number of groups resulting from the automorphisms of this section is 25.

6. *Groups of order* 64 *which contain an abelian subgroup of order* 32 *but no operator of order* 8 *which appears in such a subgroup.* If H is of type $(2, 2, 1)$ and the central is of type $(2, 2)$ there is only one group. The three abelian subgroups of index 2 contained in this group are of the same type. When the central is of type $(2, 1, 1)$ and the commutator subgroup is the characteristic subgroup of order 2 contained in the central, there are two groups. When this commutator subgroup is the square of an operator of order 4 in H which is not in the central, there are three groups, and when it is another subgroup of order 2 contained in H, there are two additional groups. Hence there are eight groups of order 64 which contain the abelian subgroup of type $(2, 2, 1)$ and two other abelian subgroups which do not involve an operator of order 8.

When the commutator subgroup is the non-cyclic group of order 4, the central must include the characteristic subgroup of this order contained in H. There are three such groups in which every operator of H corresponds to its inverse. If the commutator subgroup and the central remain the same, but only half of the operators of H correspond to their inverses, there are again three groups, while there are two additional groups when none of the operators corresponds to its inverse. When the central remains the same, but the non-cyclic commutator subgroup of order 4 is changed, there are three groups when half the operators of H correspond to their inverses, and there are two groups when this is not the case. When the central of G is of type $(2, 1)$ and

the commutator subgroup is the non-cyclic group of order 4, there are three groups in which half the operators of H correspond to their inverses, and two groups arise from the other possible automorphism.

A cyclic commutator subgroup of order 4 implies that the central is of type $(2, 1)$, and hence it results that there are two such groups. Therefore there are twenty groups of order 64 which involve the abelian group of type $(2, 2, 1)$ and give rise to a commutator subgroup of order 4 with respect to this abelian group. Suppose that in an abelian group of order 2^m an operator of order 2^a gives rise in an automorphism of order 2 to a commutator of order 2^a. The square of this operator is multiplied in this automorphism by an operator whose order is equal to this square. The central co-set to which this square belongs must involve an operator whose order is twice the order of this square since every commutator which arises from an automorphism of order 2 corresponds to its inverse in this automorphism. This proves the following theorem: *If an operator of order 2^a of an abelian group of order 2^m gives rise to a commutator of order 2^a in an automorphism of order 2, then the central under this automorphism must involve an operator of order 2^a which generates the same operator of order 2 as the commutator generates.* From this theorem it results directly that an automorphism of order 2 of the abelian group of type $(2, 2, 1)$ cannot give rise to a commutator subgroup of order 8 since this commutator subgroup would be of type $(2, 1)$.

If a group of order 64 involves the abelian subgroup of type $(2, 1, 1, 1)$ but no abelian subgroup of index 2 which has less than four independent generators, and its central is of type $(1, 1, 1, 1)$, there are three groups when every operator of H is transformed into its inverse and there is one group when the central is of type $(2, 1, 1)$. When the central is of type $(1, 1, 1, 1)$ but not every operator of an abelian subgroup of index 2 corresponds to its inverse, there are two more groups, so that there are six groups which have a commutator subgroup of order 2. If the order of the commutator subgroup exceeds 2, this subgroup cannot include any operator of order 4 and hence its order cannot exceed 4 in accord with the theorem noted in the preceding paragraph. If the central is of type $(1, 1, 1)$ and the commutator subgroup involves the characteristic subgroup of order 2, there are three groups when half of the operators of order 4 correspond to their inverses, and there are two groups in which none of these operators corresponds to its inverse.

When the central remains of type $(1, 1, 1)$, but the commutator subgroup does not involve the characteristic subgroup of order 2, there are two groups. When the central is of type $(2, 1)$ there are again two groups. Hence there are fifteen groups of order 64 which involve the abelian group of type

$(2, 1, 1, 1)$ but no abelian subgroup of index 2 which has less than four independent generators. When all the abelian subgroups of index 2 in G are of type $(1, 1, 1, 1, 1)$, the commutator subgroup must be of order 4 and there are two groups. In one of these, all the additional operators are of order 4, while only three-fourths of them have this property in the other. The total number of non-abelian groups of order 64 which involve an abelian subgroup of index 2 is therefore 118.

GROUPS CONTAINING NO ABELIAN SUBGROUP OF INDEX 2.

It was noted in § 1 that every group of order 64 contains an invariant abelian subgroup of order 16. The groups which remain to be considered may be divided into three distinct categories. The first of these is composed of all those which contain only one abelian subgroup of order 16. The second comprises all those which involve more than one such abelian subgroup but no two of them appear in a subgroup of order 32. The third category is composed of those which involve at least one subgroup of order 32 which contains three abelian subgroups of order 16. If such a group of order 64 contains only one abelian subgroup of order 16, this subgroup cannot be cyclic since such a subgroup must be transformed by the rest of the operators of the group according to a subgroup of order 4 in its group of isomorphisms, and every subgroup of order 4 in the group of isomorphisms of the cyclic group of order 2^m, $m > 2$, contains an operator of order 2 which transforms one-half of the operators of this cyclic group into themselves.

The group of isomorphisms of the abelian group of order 16 and of type $(3, 1)$ can clearly be represented as a transitive group of degree 8 with respect to letters corresponding to its operators of order 8, and when it is thus represented it contains a subgroup of order 8 formed by a $(2, 2)$ isomorphism between two regular non-cyclic groups of order 4. The remaining operators of this transitive group interchange the two cyclic subgroups of order 8 and are composed of four operators of order 2 and four of order 4. As the squares of the latter correspond to the automorphism in which every operator corresponds to its fifth power, it results that a group of order 64 which involves no abelian subgroup of order 16 besides the one of type $(3, 1)$ cannot transform the operators of this subgroup according to a cyclic group of order 4. Two of the operators of order 2 in the group of isomorphisms of this abelian group which interchange its cyclic subgroups of order 8 are commutative with all of its operators which are not of highest order, and this is also the case of the product of the remaining two operators of order 2 which interchange these cyclic subgroups. It therefore results that if a group of order 64 con-

tains only one abelian subgroup of order 16, this subgroup cannot be of type $(3, 1)$.

7. *Groups containing only one abelian subgroup of order* 16. In what precedes we used the symbol H to represent an abelian subgroup of index 2 but in this and the next section it will be convenient to use this symbol to represent the single abelian subgroup of order 16 contained in G since the properties of this subgroup will be frequently under consideration. From what precedes, it results directly that H must be one of the three groups whose types are $(1, 1, 1, 1)$, $(2, 1, 1)$, $(2, 2)$. We shall first determine all the possible groups when H is of type $(1, 1, 1, 1)$, and hence the group of isomorphisms of H is the alternating group of degree 8. This group involves two sets of conjugate operators of order 2. Under one of these half of the operators of H correspond to themselves and this operator corresponds to the substitution of degree 8 in the said alternating group. Hence it cannot appear in the group of order 4 according to which H is transformed under G. As this operator appears in one of the two sets of conjugate cyclic subgroups of order 4 contained in the group of isomorphisms of H it results that H appears in only one group of order 64 which transforms the operators of H according to a cyclic group of order 4.

There are two sets of conjugate non-cyclic groups of order 4 in the group of isomorphisms of H which correspond to subgroups of the alternating group which involve no substitution of degree 8. One of these sets corresponds to regular groups, and each group is composed of operators which are commutative with all the operators of the same subgroup of order 4 contained in H. Hence the operators of G cannot transform the operators of H according to this subgroup of order 4 since the resulting group would involve more than one abelian subgroup of order 16. The operators of G must therefore transform the operators of H according to a group of order 4 which corresponds to the non-cyclic group of order 4 and of degree 6 contained in the alternating group of degree 8. Hence there is also only one group of order 64 which transforms H according to a non-cyclic group of order 4, and there are two groups of order 64 which contain this H but no other abelian subgroup of order 16.

When H is of type $(2, 1, 1)$ the Sylow subgroup of order 64 in its group of isomorphisms is the same as before, but this group of isomorphisms contains more sets of conjugate operators whose orders are powers of 2 than the preceding one. In particular, there are now three sets of conjugate cyclic subgroups of order 4 but only one of these involves operators of order 2 which are not commutative with half the operators of H. Hence there is only one

group of order 64 which contains this H and transforms it according to a cyclic group of order 4. To prove that there is also only one set of conjugate non-cyclic subgroups of order 4 which give rise to a group of order 64 which satisfies the given conditions, it may be noted that when the group of isomorphisms of this H is represented as a subgroup of order 192 in the holomorph of the regular group of order 8 and of type $(1, 1, 1)$, the subgroup of order 24 composed of all the substitutions which omit one letter of this group of order 192 is transitive on six letters. It contains two sets of conjugate non-cyclic subgroups of order 4. These are of degree 4 and 6 respectively. As only the latter can be used according to the considerations noted above it results that there are only two groups of order 64 which involve the abelian group of type $(2, 1, 1)$ but no other abelian group of order 16.

If H is of type $(2, 2)$ the Sylow subgroup whose order is a power of 2 in its group of isomorphisms is of order 32 and contains 19 operators of order 2 and 12 operators of order 4. It can be represented as a transitive group of degree 8 involving the direct product of two regular non-cyclic groups of order four. One of these corresponds to a cyclic commutator subgroup of order 4 and hence it requires no consideration here. The other two correspond to commutator subgroups of order 8. In the square of one of these all the operators of H correspond to their inverses while only half of them correspond to their inverses in the square of the other. As one group corresponds to each of these automorphisms there are two groups of order 64 which transform this H according to a cyclic group of order 4 and contain only one abelian subgroup of order 16.

When H is transformed according to a non-cyclic group of order 4 the commutator subgroups arising from its operators of order 2 cannot all be non-cyclic. Hence two of them must be cyclic groups of order 4 while the third is a non-cyclic group of this order. There are two such non-conjugate subgroups of order 4 in the group of isomorphisms of H. In one of these the operator of order 2 which gives rise to a non-cyclic commutator subgroup transforms every operator of H into its inverse while in the other it transforms only half of these operators into their inverses. In the former case there are two groups of order 64 while there is only one such group in the latter case. Hence there are five groups of order 64 which involve this H but no other abelian subgroup of order 16. Two of these transform it according to a cyclic group of order 4 while the other three transform it according to a non-cyclic group of this order. The total number of groups of order 64 which contain only one abelian subgroup of order 16 but no abelian subgroup of order 32 is therefore 9.

8. *Groups containing five abelian subgroups of order* 16 *and are generated by every pair of them.* When G contains at least two abelian subgroups of order 16 and is generated by every pair of its abelian subgroups of this order then H has a subgroup of order 4 in common with every other abelian subgroup of order 16. Let H_1 be one such subgroup and let K represent the cross-cut of H and H_1. It is easy to see that K is non-cyclic and that every operator of G which is not also in K has four conjugates under G. The central quotient group G/K involves only operators of order 2 and hence it is abelian. Every operator of G that is not also in K appears in one and only one abelian subgroup of order 16 and hence G must involve five invariant abelian subgroups of this order, and it is generated by every pair of these subgroups. Hence we have established the following theorem: *If a group of order* 64 *contains no abelian subgroup of order* 32 *but involves at least two abelian subgroups of order* 16 *and is generated by every pair of such subgroups contained therein then it contains exactly five abelian subgroups of order* 16,[1] *and each of these subgroups is invariant.*

Suppose that at least one of the two abelian subgroups of order 16 contained in G is of type $(1,1,1,1)$. Exactly one-fourth of the remaining operators of G must be of order 2 since every operator of order 4 in G is transformed into its inverse by one-fourth of the operators of this subgroup. It therefore results that when G contains one abelian subgroup of type $(1,1,1,1)$ it also contains an invariant abelian subgroup H of type $(2,2)$. If an operator of the abelian subgroup of type $(1,1,1,1)$ transforms into their inverses all the operators of this H the remaining operators of the former subgroup must transform H in two ways determined by the two possible automorphisms of order 3 of the four group. On the other hand if half of the operators of H are transformed into their inverses by an operator of the said group of type $(1,1,1,1)$ the other two transformations under this subgroup are again completely determined, and hence there are two possible such groups of order 64 which involve an abelian subgroup of order 16 and of type $(1,1,1,1)$.

It remains to determine the groups in which the five abelian subgroups of order 64 are of the types $(2,2)$, $(2,1,1)$. There is only one such group in which each of the five abelian subgroups is of type $(2,2)$ and there is no such group in which each of these subgroups is of type $(2,1,1)$. Hence it results that *every group of order* 64 *which involves five abelian subgroups of order* 16 *but no subgroup of order* 32 *which is either abelian or has a central of order* 8 *contains at least one abelian subgroup of type* $(2,2)$. There is one such group which involves four abelian subgroups of type $(2,1,1)$ and hence twenty operators of order 2, and there is one which contains three

abelian groups of type $(2, 2)$ and two of type $(2, 1, 1)$, and hence twelve operators of order 2. The total number of groups of order 64 which contain no abelian subgroup of order 32 but five abelian subgroups of order 16, and are generated by every pair of these subgroups is therefore 5. It remains to determine the groups of order 64 which involve a non-abelian subgroup of index 2 containing three abelian subgroups of order 16.

9. *Groups of order* 64 *arising from three-fourths automorphisms.* In what follows the symbol H will be used to represent a non-abelian subgroup of index 2 which contains three abelian subgroups of order 16. It is known that a characteristic property of such a group is that it admits three automorphisms in which exactly three-fourths of the operators correspond to their inverses. Such automorphisms have been called *three-fourths automorphisms,** and in the present section we aim to determine all the groups of order 64 which arise from a three-fourths automorphism of H but do not involve an abelian subgroup of order 32. Since one of the products of a three-fourths automorphism and the group of inner automorphisms is a characteristic operator in the group of automorphisms and the operators of the central correspond to their inverses under this automorphism, it results that the multiplier s_0 in the general theory of constructing all the groups which contain a given group as an invariant subgroup of prime index † can be selected only from the different sets of conjugates of order 2 in the central of H. Hence the following theorem: *The number of the distinct groups which involve as a subgroup of index* 2 *a given group* H *and transform it according to a three-fourths automorphism is equal to one more than the number of the different sets of operators of order* 2 *in the central of* H *which are conjugate under the group of isomorphisms of* H.

A group resulting from this theorem cannot involve an abelian subgroup of index 2 unless the central of H is of type $(1, 1, 1 \cdots)$. Moreover, when one of the abelian subgroups of index 2 in H is of type $(1, 1, 1 \cdots)$ the resulting group must involve an abelian subgroup of index 2. It remains therefore to consider the case when each of the abelian subgroups of index 2 in H is of type $(2, 1, 1 \cdots)$. If the resulting group does not contain an abelian subgroup of index 2 the operators of order 4 in H must have three distinct squares and the commutator of order 2 is the product of these squares. When H is of order 16 its central must therefore involve three operators of order 2 which are squares, three others which are the product of two squares

* G. A. Miller, *Proceedings of the National Academy of Sciences*, Vol. 15 (1929), p. 269.

† *Ibid.*, Vol. 14 (1928), p. 819.

and a commutator of order 2. Hence *there are four groups of order* 64 *which arise from an* H *whose central is of type* $(1, 1, 1)$ *by means of a three-fourths automorphism and do not involve an abelian subgroup of order* 32. It results directly from the theorem noted at the close of the preceding paragraph that when the central of H is a cyclic group of even order there are exactly two groups which arise from a three-fourths automorphism of H. Hence there are four such groups of order 64, two when H involves an operator of order 16 and two more when H involves three abelian subgroups of type $(3, 1)$.

When H involves an abelian subgroup of type $(3, 1)$ but its central is non-cyclic there are two such H's in which the commutator subgroup is generated by the square of an operator in the central and two others in which this is not the case. Each of the former gives rise to three groups while each of the latter gives rise to four groups, according to the general theorem noted above. Hence the total number of groups which results from these four H's when they are transformed according to a three-fourths automorphism is 14. When H involves no operator of order 8 but contains an abelian subgroup of type $(2, 2)$ there are three groups of order 32 which may be used for H. In one of these the commutator subgroup is generated by an operator of order 4 contained in the central, and this gives rise to three groups of order 64 while each of the other two gives rise to four groups of this order but only seven of these are distinct. It remains to consider the case when each of the three abelian subgroups of H is of type $(2, 1, 1)$, and the central of H is of type $(2, 1)$. Three groups of order 64 arise from this H, and hence the total number of groups of order 64 which arise from three-fourths automorphisms and involve no abelian subgroup of order 32 is 35.

10. *Groups of order* 64 *arising from other automorphisms.* It remains to consider the groups of order 64 which involve a subgroup H of index 2 admitting a three-fourths automorphism but are not transformed according to such an automorphism under G and do not contain any other subgroup of this index which is either abelian or thus transformed. When H contains a cyclic subgroup of order 16 all the operators of H may be transformed into their fifth powers under G and there is one such group and there is also one such group when the operators of one cyclic subgroup of order 16 are transformed into their third powers while those of the other are transformed into their eleventh powers. There are four sets of four operators with respect to the group of inner automorphisms which interchange the two cyclic groups of order 16 in H. In one of these sets there is an operator which is commutative with all the operators of an abelian subgroup of order 16 in H and hence only three of these sets require consideration here. The one which

involves only operators of order 4 gives rise to two groups while each of the others gives rise to only one group. There are therefore six additional groups when H contains an operator of order 16.

When H involves three abelian subgroups of type $(3, 1)$ and its central is the cyclic group of order 8 there is one automorphism of order 2 which transforms the operators of highest order in the central and in another cyclic group of order 8 into their fifth powers but is commutative with the remaining eight operators of this order. There is one group which corresponds to this automorphism. Under the Sylow subgroup of order 2^m in the group of isomorphisms of H these remaining operators of order 8 are transformed according to a transitive group of degree 8 and order 16 and hence this Sylow subgroup is the direct product of this group of order 16 and a group of order 2. There is clearly only one group which transforms H according to the product of a three-fourths automorphism and the given invariant operator of order 2. It remains to consider the cases when two cyclic groups of order 8 are transformed into each other. Four of these operators would give rise to an abelian subgroup of order 32 and hence they do not require consideration here. The product of one of these and the said invariant operator of order 2 gives rise to one group while the product of this and an operator which transforms H into a three-fourths automorphism gives rise to two groups. As the remaining product gives rise to only one group there are six groups which arise from this H.

There are four other H's which involve an abelian subgroup of type $(3, 1)$. When the commutator subgroup of H is generated by an operator of order 4 contained in such an H and one of the abelian subgroups of H is of type $(2, 1, 1)$ there are sixteen operators in the group of isomorphisms of H which transform each of its cyclic subgroups of order 8 into itself and these four subgroups are transformed according to the octic group under the group of isomorphisms. The said 16 operators are generated by a three-fourths automorphism and the 8 operators which leave invariant separately each operator in a subgroup of index 2 contained in H. These 8 operators do not give rise to a new group. Hence we need to consider only one of the operators which transform into itself each of the four cyclic groups of order 8 contained in H and this gives rise to only one group.

There are eight operators which interchange two of the four cyclic subgroups of order 8 which do not give rise to an additional group. The remaining eight give rise to six groups,—one set of four giving rise to four groups while the other gives rise to only two groups. When the four cyclic subgroups of order 8 are interchanged according to the invariant operator of order 2 in the octic group there is only one set of four operators which

do not give rise to an additional group while there are eight groups which correspond to the remaining three such sets. When these subgroups are transformed according to a non-invariant substitution of order 2 and degree 4 in the octic group there is again a set of four operators which does not give rise to an additional group, while the other three sets of four operators give rise to seven groups. As there is obviously no group which transforms the four cyclic groups of order 4 in H according to an operator of order 4 in the given octic group this H gives rise to 22 additional groups.

We proceed to consider the possible groups which result from the H which has the same commutator subgroup as the preceding H but one of the abelian subgroups is of type $(2, 2)$ instead of type $(2, 1, 1)$. The four cyclic subgroups of order 4 in this H are again transformed according to the octic group under its group of isomorphisms but the number of subgroups of index 2 in this H is 3 while there are 7 such subgroups in the preceding H. Hence there are only eight operators in the group of isomorphisms of the present H which transform into itself each of its cyclic subgroups of order 8. Hence there are 3 groups of order 64 which contain this H and interchange only two of its cyclic subgroups of order 8 and there are 3 more which transform the four cyclic subgroups in pairs according to the invariant operator of order 2 in the octic group. As there are four groups when these four subgroups are transformed according to a non-invariant operator in the octic group, the number of the groups which result from this H is 10.

When H has as a commutator subgroup a non-characteristic subgroup of order 2 its group of isomorphisms is again of order 64 and there are four groups which interchange only two of its cyclic subgroups of order 8. There are also four groups which interchange these cyclic subgroups according to the invariant operator of order 2 in the octic group. As there are also four groups when the four subgroups are permuted according to a non-invariant operator of order 2 in the octic group the number of groups which arise from this H is 12. There is only one other possible H which involves operators of order 8 and there are again only 8 operators in its group of isomorphisms which transform each of the four cyclic subgroups of order 8 in H into itself. Just as in the preceding case there result 8 groups when only two cyclic subgroups of order 8 in H are interchanged or when these cyclic subgroups are interchanged according to the invariant operator of order 2 in the octic group. When the four cyclic subgroups of order 8 are interchanged according to a non-invariant operator of order 2 in the octic group there are also four groups, and hence the number of groups which result from this H is also 12. The total number of groups which involve an H containing operators of order 8 but none of order 16 is therefore 62.

It remains to consider the groups which do not involve an H containing operators of order 8. There is one such H which involves three abelian subgroups of type $(2, 2)$ and its group of isomorphisms is of order 32. Half of these transform into itself all the operators of the central of H and include the group of inner isomorphisms but no three-fourths isomorphism. These give rise to groups which contain an abelian subgroup of index 2 and hence they need no consideration here. Hence there are only four groups resulting from this H. When two of the abelian subgroups are of type $(2, 2)$ and the third is of type $(2, 1, 1)$ there are six groups which arise from the automorphisms which are similar to those of the preceding H. The present H has, however, 32 additional automorphisms which transform into each other two operators of order 2 in its central. There are twelve groups which correspond to these automorphisms. When only one of the abelian subgroups of order 16 in H is of type $(2, 2)$ its group of isomorphisms is of order 32 and the three operators of order 2 in the central of H are characteristic. Eight groups result from this H. When the three abelian subgroups of order 16 are of type $(2, 1, 1)$ and the central is of type $(2, 1)$ the group of isomorphisms of H contains a Sylow subgroup of order 64, and H involves 15 subgroups of index 2. The groups arising from these isomorphisms are covered by the following theorem: *There are four and only four groups of order* 64 *which satisfy the following conditions: They involve no operator whose order exceeds* 4 *and their operators of order* 4 *have a common square which generates their commutator subgroup.* Two of these four groups arise from three-fourths automorphisms while the other two do not have this property. The latter contain the H under consideration.

When half of the operators of H not including the central are transformed into themselves and the rest into themselves multiplied by a non-characteristic operator of order 2 there are two groups. When all of the abelian subgroups of index 2 in H are of type $(2, 1, 1)$ and the central is of type $(1, 1, 1)$ there are two groups when the operators of order 4 have a common square and eight groups when the squares are not the same. When one of the abelian subgroups is of type $(1, 1, 1, 1)$ and the other two are of type $(2, 1, 1)$ there are four groups. Finally, when two of the abelian subgroups of order 16 are of type $(1, 1, 1, 1)$ there are no additional groups. Hence the total number of groups of order 64 which do not contain a subgroup of order 32 involving an abelian subgroup of order 16 which contains operators of order 8 is 48 and the total number of groups of order 64 is 294. The number of those which do not involve an abelian subgroup of index 2 is 165.

A General System of Ordinary Differential Equations of the First Order.*

By Eugene Feenberg.

Introduction to Part 1.

The object of this section is to determine necessary and sufficient conditions that a positive valued scalar function $H[w_1(x), w_2(x), \cdots, w_j(x) \cdots]$, satisfy the inequality

$$1.1 \quad H[\int_a^x w_1(t)\,dt, \int_a^x w_2(t)\,dt, \cdots] \leqq \int_a^x H[w_1(t), w_2(t), \cdots]\,dt.$$

The following notation and definitions are used thruout the paper: Vectors are represented by Clarendon type, scalars and components of vectors by light face type. $\mathbf{w} \equiv (w_1, w_2, \cdots)$, an ordered set of numbers is a vector. $\mathbf{w}(k)$ is the vector obtained from $\mathbf{w}$ by taking

$$\begin{aligned} w_j(k) &= w_j, \qquad && j \leqq k, \\ w_j(k) &= 0, && j > k. \end{aligned}$$

A vector is said to have a certain property when all the components of the vector have that property. A scalar function of a vector, $F(\mathbf{w})$, will be written as $F(w_j)$ when all the components of $\mathbf{w}$ except w_j are zero. A vector is on the "extended" range $\mathbf{D}$, [$\mathbf{D}$ a sequence of positive numbers, $(D_1, D_2, \cdots)$], if there exists a positive number e such that $-eD_j < w_j < eD_j$, $(j = 1, 2, \cdots)$.

Part 1.

Consider an arbitrary range $\mathbf{D}$ and a positive valued scalar function, $H(\mathbf{w})$, defined for all values of $\mathbf{w}$ on the extended range $\mathbf{D}$. In the following theorems the symbols $\mathbf{w}, \mathbf{w}', \mathbf{w}'', \mathbf{z}, \mathbf{z}', \mathbf{z}''$ represent vectors on the extended range $\mathbf{D}$.

Theorem I. *Necessary conditions that $H(\mathbf{w})$ satisfy the inequality $H[\int_a^x \mathbf{w}(t)\,dt] \leqq \int_a^x H[\mathbf{w}(t)]\,dt$, where $\mathbf{w}(x)$ is a summable vector, $H[\mathbf{w}(x)]$ is summable, and $\int_a^x \mathbf{w}(t)\,dt$ is on the extended range $\mathbf{D}$, $(a \leqq x \leqq b)$, are*

* Presented to the American Mathematical Society, September 7, 1928.

$$H(cw)=cH(w), \quad c\geqq 0, \qquad H(w'+w'')\leqq H(w')+H(w'').$$

Proof. Take $w(x)=w$, (w a constant vector). Then $H(\int_0^x w\,dt) = H(xw)$ and $\int_0^x H(w)\,dt = xH(w)$. By (1.1)

1.2 $$H(cw)\leqq cH(w), \quad c\geqq 0.$$

c and w are arbitrary; replacing c by $1/c$ and w by cw $H(w)\leqq 1/cH(cw)$.

1.3 $$H(cw)\geqq cH(w).$$

Hence

1.4 $$H(cw)=cH(w), \quad c\geqq 0.$$

Take $w(x)=w'$, $0\leqq x\leqq 1$, $w(x)=w''$, $1<x\leqq 2$. By (1.1)

$$H(\int_0^1 w'dt+\int_1^2 w''dt)\leqq \int_0^1 H(w')\,dt+\int_1^2 H(w'')\,dt.$$

Hence

1.5 $$H(w'+w'')\leqq H(w')+H(w'').$$

It will be shown that conditions (1.4) and (1.5) are sufficient to prove the inequality (1.1) for the case of a finite number of variables.

THEOREM II. (1.4) *and* (1.5) *are sufficient conditions that* $H[w(k)]$ *be continuous in* $w(k)$, $(k=1,2,\cdots)$.

Proof. Define I_j as the vector with every component zero except the j-th component which is identically one. By (1.4)

$$H(cw)=cH(w), \quad H[c(-w)]=cH(-w), \qquad c\geqq 0.$$

Take $w=I_j$, $c=|v_j|$, v_j an arbitrary number. Then

$$H(cw)=H(|v_j|I_j)=|v_j|H(I_j).$$

Take $w=-I_j$, $c=|v_j|$.

$$H(cw)=H(-|v_j|I_j)=|v_j|H(-I_j).$$

Hence

1.6 $$H(w_j)\leqq |w_j|[H(I_j)+H(-I_j)],$$

Applying (1.5) k times

1.7 $$H[w(k)]\leqq \sum_{j=1}^{k} H(w_j)\leqq \sum_{j=1}^{k}|w_j|[H(I_j)+H(-I_j)], \qquad (k=1,2,\cdots).$$

Thus since $H(0)=0$ and $\underset{w(k)\to 0}{\text{Limit}}\, H[\boldsymbol{w}(k)]=0$, $H[\boldsymbol{w}(k)]$ is continuous at $\boldsymbol{w}(k)=0$, $(k=1,2,\cdots)$, Let $\boldsymbol{w}'+\boldsymbol{w}''=\boldsymbol{z}'$, $\boldsymbol{w}''=\boldsymbol{z}''$. By (1.5)

1.8 $$H(\boldsymbol{z}'-\boldsymbol{z}'')\geqq H(\boldsymbol{z}')-H(\boldsymbol{z}'')$$

and similarly

1.9 $$H(\boldsymbol{z}''-\boldsymbol{z}')\geqq H(\boldsymbol{z}'')-H(\boldsymbol{z}').$$

Take $\boldsymbol{z}'=\boldsymbol{z}'(k)$, $\boldsymbol{z}''=\boldsymbol{z}''(k)$. Then

$$\underset{z''\to z'}{\text{Limit}}\, H[\boldsymbol{z}'(k)-\boldsymbol{z}''(k)]=\underset{z''\to z'}{\text{Limit}}\, H[\boldsymbol{z}''(k)-\boldsymbol{z}'(k)]=0,$$
$$\underset{z''\to z'}{\text{Limit}}\, \{H[\boldsymbol{z}'(k)]-H[\boldsymbol{z}''(k)]\}\leqq 0,$$
$$\underset{z''\to z'}{\text{Limit}}\, \{H[\boldsymbol{z}''(k)]-H[\boldsymbol{z}'(k)]\}\leqq 0.$$

Hence $\underset{z''\to z'}{\text{Limit}}\, H[\boldsymbol{z}''(k)]=H[\boldsymbol{z}'(k)]$ and $H[\boldsymbol{w}(k)]$ is continuous in $\boldsymbol{w}(k)$, $(k=1,2,\cdots)$.

Take $(\boldsymbol{w}^{(1)},\boldsymbol{w}^{(2)},\cdots)$ a sequence of vectors such that each vector and $\sum_{j=1}^{\infty}\boldsymbol{w}^{(j)}$ are on the extended range $\boldsymbol{D}$ and $\sum_{j=1}^{\infty}H(\boldsymbol{w}^{(j)})$ converges. That such a sequence exists is readily shown from the inequality (1.7). Define

$$\boldsymbol{w}(x)=2^{j}\boldsymbol{w}^{(j)},\quad 2^{-j}<x\leqq 2^{-j+1},\qquad (j=1,2,\cdots).$$

Then by (1.1)

1.10 $$H[\sum_{n+1}^{\infty}\boldsymbol{w}^{(j)}]=H[\int_0^{2^{-n}}\boldsymbol{w}(t)\,dt]\leqq\int_0^{2^{-n}}H[\boldsymbol{w}(t)]\,dt=\sum_{n+1}^{\infty}H(\boldsymbol{w}^{(j)}).$$

Hence

1.11 $$\underset{n\to\infty}{\text{Limit}}\, H[\sum_{j=n}^{\infty}\boldsymbol{w}^{(j)}]=0.$$

In particular if $\boldsymbol{w}$ is a vector such that $\sum_{j=n}^{\infty}[H(w_j)+H(-w_j)]$ exists, then $\underset{k\to\infty}{\text{Limit}}\, H[\boldsymbol{w}-\boldsymbol{w}(k)]=\underset{k\to\infty}{\text{Limit}}\, H[\boldsymbol{w}(k)-\boldsymbol{w}]=0$, and $\underset{k\to\infty}{\text{Limit}}\, H[\boldsymbol{w}(k)]=H(\boldsymbol{w})$.

To prove the sufficiency of conditions (1.4) and (1.5) we make use of a corollary to Valleé Poussin's * integral theorem. For clearness the theorem is restated here in a restricted form.

Given a sequence of functions $f_n(x)\geqq 0$, $(n=1,2,\cdots)$, and a function $f(x)\geqq 0$ such that on the interval (a,b)

* "Sur L'Intégrale de Lebesgue," *Transactions of the American Mathematical Society*, Vol. 16 (1915), p. 435.

c_1. $\underset{n\to\infty}{\text{Limit}}\, f_n(x) = f(x)$ "a. e.",

c_2. $f_n(x)$ is summable, $(n = 1, 2, \cdots)$,

c_3. $\int_a^x f_n(t)\,dt$ is absolutely continuous uniformly with resp

Conclusion: c_1, c_2, c_3 are necessary and sufficient for c_4.

c_4. $f(x)$ is summable and $\underset{n\to\infty}{\text{Limit}} \int_a^x f_n(t)\,dt = \int_a^x f(t)\,dt.$

interval (a, b) into n subintervals taking $a = t_0 \leqq t_{1n} \cdots \leqq t_{nn}$ points of division. A horizontal function of index n, $h_n(x)$, is d a function such that $h_n(x) = h_{in}$, (h_{in} a constant), on $t_{in} \leqq$ If $F(x)$ is a summable function on (a, b) there exist for any me subdivision, a sequence of horizontal functions, $h_n(x)$, such that

$= F(x)$ "a. e." on (a, b), $\underset{n\to\infty}{\text{Limit}} \int_a^x h_n(t)\,dt$ and $\underset{n\to\infty}{\text{Limit}} \int_a^x | h_n($ independent of the method of fine subdivision. Take $F(x)$ a summable function and $h_n(x) \geqq 0$. Then the corollary states that is absolutely continuous in x uniformly with respect to n. To identify $h_n(x)$ with $f_n(x)$ and $F(x)$ with $f(x)$. c_1, c_2, c_4 of stated above are satisfied. Therefore c_3 holds proving the coroll

THEOREM III. *Given the conditions* (1.4), (1.5) *and a vec summable on* (a, b), *then*

$$H[\int_a^x \mathbf{g}(t, k)\,dt] \leqq \int_a^x H[\mathbf{g}(t, k)]\,dt, \ (a \leqq x \leqq b), \quad (k =$$

Proof. There exists a sequence of horizontal vectors, $\mathbf{h}^{(n)}$ that $\underset{n\to\infty}{\text{Limit}}\ h_j^{(n)}(x) = g_j(x)$ "a. e." on (a, b), $\underset{n\to\infty}{\text{Limit}} \int_a^x h$ $\int_a^x g_j(t)\,dt$ and $\underset{n\to\infty}{\text{Limit}} \int_a^x | h_j^{(n)}(t) |\,dt = \int_a^x | g_j(t) |\,dt$, $(j =$ By theorem II.

$$\underset{n\to\infty}{\text{Limit}}\ H[\mathbf{h}^{(n)}(x, k)] = H[\mathbf{g}(x, k)] \text{ "a. e." on } (a, b).$$

$$\underset{n\to\infty}{\text{Limit}}\ H[\sum_{i=1}^{i_{xn}} \mathbf{h}^{in}(k)\Delta_{in}] = \underset{n\to\infty}{\text{Limit}}\ H[\int_a^x \mathbf{h}^{in}(t, k)\,dt] = H[\int_a^x$$

By conditions (1.4) and (1.5)

---

* H. J. Ettlinger, "On Multiple Iterated Integrals," *American Journ matics*, Vol. 48 (1926), p. 215.

† i_{xn} is for each value of n the subscript of the last division point point x, i. e., $t_{i_{xn}} < x \leqq t_{i_{xn+1}}$.

$$H[\sum_{i=1}^{i_{xn}} \boldsymbol{h}^{in}(k)\Delta_{in}] \leqq \sum_{i=1}^{i_{xn}} H[\boldsymbol{h}^{in}(k)]\Delta_{in}.$$

1.12 $$\int_{x_1}^{x_2} H[\boldsymbol{h}^{(n)}(t,k)]dt \leqq \int_{x_1}^{x_2} \sum_{j=1}^{k} H[h_j^{(n)}(t)]dt$$
$$\leqq \sum_{j=1}^{k} [H(I_j)+H(-I_j)] \int_{x_1}^{x_2} |\, h_j^{(n)}(t) \,|\, dt.$$

Therefore

$$\operatorname*{Limit}_{n\to\infty} \{ \sum_{i=1}^{i_{xn}} H[h^{in}(k)]\Delta_{in} - \int_a^x H[\boldsymbol{h}^{(n)}(t,k)]dt\} = 0, \int_a^x H[\boldsymbol{h}^{(n)}(t,k)]dt$$

is absolutely continuous in x uniformly with respect to n and

$$\operatorname*{Limit}_{n\to\infty} \sum_{i=1}^{i_{xn}} H[\boldsymbol{h}^{in}(k)]\Delta_{in} = \int_a^x H[\boldsymbol{g}(t,k)]dt.$$

Hence

1.13 $$H[\int_a^x \boldsymbol{g}(t,k)dt] \leqq \int_a^x H[\boldsymbol{g}(t,k)]dt, \ (k=1,2,\cdots).$$

THEOREM IV. *Given conditions* (1.4), (1.5) and a summable vector, $\boldsymbol{g}(x)$, *such that* $H[\boldsymbol{g}(x)]$ *is finite and* $\operatorname*{Limit}_{k\to\infty} H[\boldsymbol{g}(x,k)] = H[\boldsymbol{g}(x)]$ "a. e." *on* (a,b), $\operatorname*{Limit}_{k\to\infty} H[\int_a^x \boldsymbol{g}(t,k)dt] = H[\int_a^x \boldsymbol{g}(t)dt]$, $\int_a^x H[\boldsymbol{g}(t,k)]dt$ *is absolutely continuous in x uniformly with respect to k. Then*

$$H[\int_a^x \boldsymbol{g}(t)dt] \leqq \int_a^x H[\boldsymbol{g}(t)]dt.$$

Proof. By theorem III.

$$\operatorname*{Limit}_{k\to\infty} H[\int_a^x \boldsymbol{g}(t,k)dt] \leqq \operatorname*{Limit}_{k\to\infty} \int_a^x H[\boldsymbol{g}(t,k)]dt.$$

By Valleé Poussin's theorem $H[\boldsymbol{g}(x)]$ is summable and

$$\operatorname*{Limit}_{k\to\infty} \int_a^x H[\boldsymbol{g}(t,k)]dt = \int_a^x H[\boldsymbol{g}(t)]dt.$$

Therefore

$$H[\int_a^x \boldsymbol{g}(t)dt] \leqq \int_a^x H[\boldsymbol{g}(t)]dt.$$

. . . .

The function $H(\boldsymbol{w}) = (\sum_{j=1}^{\infty} |\, w_j \,|^{p})^{1/p}$, $p \geqq 1$, obviously has the property (1.4). That (1.5) is a property of this function has been shown by Friedrich Riesz.*

* "Untersuchungen über Systeme Integrierbarer Funktionen," *Mathematische Annalen*, Vol. 68 (1910), p. 455.

Using Riesz's result the function

$$H(\boldsymbol{w}) = [\sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} |w_j|^{q_i})^{p/q_i} c_i]^{1/p}, \qquad (p \geqq 1),\ (q_i \geqq 1),\ (c_i \geqq 0),$$

is readily found to satisfy conditions (1.4) and (1.5).

If $H(w)$ is an even function, [i. e., $H(\boldsymbol{w}) = H(-\boldsymbol{w})$], the properties

1.14 $$H(\boldsymbol{w} - \boldsymbol{w}') \geqq |H(\boldsymbol{w}) - H(\boldsymbol{w}')|,$$

1.15 $$H(\boldsymbol{w} - \boldsymbol{w}') \geqq |H(\boldsymbol{w} + \boldsymbol{z}) - H(\boldsymbol{w}' + \boldsymbol{z})|.$$

follow immediately from (1.8) and (1.9).

Introduction to Part 2.

Systems of ordinary differential and difference equations in a countable number of variables are treated in Part 2. Such systems have been studied by a number of authors.* It is the purpose of this paper to apply the results of the preceding section to the establishment of certain general existence theorems.

The set of hypotheses Q are as follows:

I. $\boldsymbol{a} \equiv (a_1, a_2, \cdots)$ is a constant vector. $\boldsymbol{D}^*(x)$ is a positive vector function bounded by a constant positive vector $\boldsymbol{R}$, i. e., $0 \leqq D_j^*(x) \leqq r_j$. The vector arguments $\boldsymbol{w}(x)$, $\boldsymbol{w}'(x)$, $\boldsymbol{w}^{(l)}(x)$, $\boldsymbol{w}'^{(l)}(x)$, $(l = 1, 2, \cdots)$, used in the following hypotheses are any set of vectors such that for any vector, $\boldsymbol{w}(x)$, of the set, $|w_j(x) - a_j| \leqq D_j^*(x) \leqq r_j$, $(j = 1, 2, \cdots)$. The properties stated are with respect to the interval $0 \leqq x \leqq b$.

II. $\boldsymbol{f}[x, \boldsymbol{w}(x)] \equiv \{f_1[x, \boldsymbol{w}(x)], f_2[x, \boldsymbol{w}(x)], \cdots\}$ is a summable vector such that $|\int_a^x f_j[t, \boldsymbol{w}(t)]dt| \leqq D_j^*(x)$, $(j = 1, 2, \cdots)$.

III. $[H_1(\boldsymbol{w}), H_2(\boldsymbol{w}), \cdots]$ is a sequence of positive valued scalar functions such that on the range $\boldsymbol{D}$: $D_j = 2(|a_j| + r_j)$, $(j = 1, 2, \cdots)$,

a. $H_j(\boldsymbol{w}) \leqq M_j$, M_j a positive number,

b. $H_j[\int_0^x \boldsymbol{g}(t)dt] \leqq \int_0^x H_j[\boldsymbol{g}(t)]dt$ when $H_j[\boldsymbol{g}(x)]$, is summable and $\int_0^x \boldsymbol{g}(t)dt$ is on $\boldsymbol{D}$, $(j = 1, 2, \cdots)$.

IV. There exist sequences of positive valued summable functions:

2.1 $$\{G_1(x, \boldsymbol{w}, \boldsymbol{w}'),\ G_2(x, \boldsymbol{w}, \boldsymbol{w}'), \cdots\},$$

---

* W. L. Hart, "Differential Equations and Implicit Functions in Infinitely Many Variables," *Transactions of the American Mathematical Society*, Vol. 18 (1917).

2.2 $\{K_1(x), K_2(x), \cdots\}$, such that

2.3 $\sum_{l=0}^{\infty} 1/l![\int_0^b K_j(t)dt]^l \int_0^x G_j[t, w^{(l)}(t), w'^{(l)}(t)]dt$ *exists.*

V. Finally hypotheses III and IV are such that:

2.4 $$|\int_{x_1}^{x_2}\{f_j[t, w(t)] - f_j[t, w'(t)]\}dt| \leqq \int_{x_1}^{x_2} G_j[t, w(t), w'(t)]H_j[w(t) - w'(t)]dt.$$

2.5 $$H_j[\int_{x_1}^{x_2}\{f[t, w(t)] - f[t, w'(t)]\}dt] \leqq \int_{x_1}^{x_2} K_j(t)H_j[w(t) - w'(t)]dt,$$
$$0 \leqq x_1 \leqq x_2 \leqq b, \qquad (j = 1, 2, \cdots).$$

Consider the system $y(x) = \int_0^x f[t, y(t)]dt + a$. A vector absolutely continuous in x and satisfying the system is a solution. It is shown that under the set of hypotheses Q the system has one and only one solution on $D^*(x)$.

A difference system is derived from the differential system and with a slight modification of the hypotheses Q the existence of a unique solution is established. Let the interval $(0, b)$ be divided into n subintervals by the points $0 = t_{0n} \leqq t_{1n} \cdots \leqq t_{nn} = b$. $w_n(x) = w_n(i)$ is defined to be a vector, the components of which are horizontal functions constant on the subintervals $t_{in} \leqq x < t_{i+1,n}$, $(i = 1, 2, \cdots, n)$. Defining

$$B_n[i-1, w_n(i-1)] \equiv \int_{t_{i-1,n}}^{t_{in}} f[t, w_n(i-1)]dt,$$

the difference system is

$$y_n(i) = \sum_{r=0}^{i-1} B_n[r, y_n(r)] + a.$$

With the additional hypotheses: $H_j[f(x, a)]$ summable, $G_j(x, w, w') = G_j(x)$, $G_j(x)$ a summable function, $(j = 1, 2, \cdots)$, the solution is shown to pass over in the limit as n becomes infinite, for any method of fine subdivision, to the solution of the corresponding differential system.

PART 2.

THEOREM V. *Given the set of hypotheses Q then the system*

$$y(x) = \int_0^x f[t, y(t)]dt + a$$

has one and only one solution on $D^*(x)$, $0 \leqq x \leqq b$.

Proof. By the usual method of successive approximation take

$$
\begin{aligned}
&\mathbf{y}^{(0)}(x)=\mathbf{a},\\
&\mathbf{y}^{(1)}(x)=\int_0^x \mathbf{f}[t,\mathbf{y}^{(0)}(t)]dt+\mathbf{a},\\
2.6\qquad &\mathbf{y}^{(2)}(x)=\int_0^x \mathbf{f}[t,\mathbf{y}^{(1)}(t)]dt+\mathbf{a},\\
&\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&\mathbf{y}^{(k)}(x)=\int_0^x \mathbf{f}[t,\mathbf{y}^{(k-1)}(t)]dt+\mathbf{a},\qquad (k=1,2,\cdots).
\end{aligned}
$$

By hypotheses Q, II. this sequence exists for all values of k. Define

$$
\begin{aligned}
&\mathbf{U}^{(1)}(x)=\mathbf{y}^{(1)}(x),\\
&\mathbf{U}^{(2)}(x)=\mathbf{y}^{(2)}(x)-\mathbf{y}^{(1)}(x),\\
2.7\qquad &\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&\mathbf{U}^{(k)}=\mathbf{y}^{(k)}(x)-\mathbf{y}^{(k-1)}(x),\qquad (k=1,2,\cdots).
\end{aligned}
$$

Then

$$\mathbf{y}^{(k)}(x)=\mathbf{y}^{(1)}(x)+\sum_{l=2}^{k}\mathbf{U}^{(l)}(x).$$

By hypotheses Q, II and Q, line 2.4

$$
2.8\qquad
\begin{aligned}
&|\,U_j^{(1)}(x)\,|\leqq r_j+|\,a_j\,|,\\
&|\,U_j^{(k)}(x)\,|\leqq\int_0^x G_j[t,\mathbf{y}^{(k-1)}(t),\mathbf{y}^{(k-2)}(t)]H_j[t,\mathbf{U}^{(k-1)}(t)]dt,\ (k\geqq 2).
\end{aligned}
$$

By hypotheses Q, III and Q, line 2.5

$$2.9\quad H_j[\mathbf{U}^{(k)}(x)]\leqq M_j,\ H_j[\mathbf{U}^{(k)}(x)]\leqq\int_0^x K_j(t)H_j[\mathbf{U}^{(k-1)}(t)]dt,\ (k\geqq 2).$$

Hence by induction

$$2.10\quad H_j[\mathbf{U}^{(k)}(x)]\leqq M_{j/(k-1)!}\Big[\int_0^b K_j(t)\,dt\Big]^{k-1},\qquad (k=1,2,\cdots),$$

$$
\begin{aligned}
2.11\quad &|\,U_j^{(k)}(x)\,|\\
&\leqq M_{j/(k-2)!}\Big[\int_0^b K_j(t)\,dt\Big]^{k-2}\int_0^b G_j[t,\mathbf{y}^{(k-1)}(t),\mathbf{y}^{(k-2)}(t)]dt,\ (k\geqq 2),
\end{aligned}
$$

$$
\begin{aligned}
2.12\quad &\sum_{l=1}^{\infty}|\,U_j^{(l)}(x)\,|\\
&\leqq\sum_{l=0}^{\infty}M_{j/l!}\Big[\int_0^b K_j(t)\,dt\Big]^{l}\int_0^b G_j[t,\mathbf{y}^{(l)}(t),\mathbf{y}^{(l-1)}(t)]dt+|\,a_j\,|+y_j.
\end{aligned}
$$

Therefore $\operatorname*{Limit}_{k\to\infty} y_j^{(k)}(x)=\operatorname*{Limit}_{k\to\infty}\sum_{l=1}^{k}U_j^{(l)}(x)$ exists [$=y_j(x)$ say].

$$2.13\qquad \operatorname*{Limit}_{k\to\infty} y_j^{(k)}(x)=y_j(x)$$

$\mathbf{D}^*(x)$ for any fixed value of x is a closed region. Therefore $\mathbf{y}(x)$ is on $\mathbf{D}^*(x)$. It remains to prove that $\mathbf{y}(x)$ is a solution of the system. By 2.6

$$\operatorname*{Limit}_{k\to\infty} y_j^{(k)}(x) = y_j(x) = \operatorname*{Limit}_{k\to\infty} \int_0^x f_j[t, \mathbf{y}^{(k-1)}(t)]dt + a_j.$$

$$\begin{aligned}
&\left| \int_0^x \{f_j[t, \mathbf{y}(t)] - f_j[t, \mathbf{y}^{(k)}(t)]\}dt \right| \\
&\leqq \int_0^x G_j[t, \mathbf{y}(t), \mathbf{y}^{(k)}(t)] H_j[\mathbf{y}(t) - \mathbf{y}^{(k)}(t)]dt \\
&= \int_0^x G_j[t, \mathbf{y}(t), \mathbf{y}^{(k)}(t)] H_j[\sum_{l=k+1}^{\infty} \mathbf{U}^{(l)}(t)]dt \\
&\leqq \int_0^x G_j[t, \mathbf{y}(t), \mathbf{y}^{(k)}(t)] \sum_{l=k+1}^{\infty} H_j[\mathbf{U}^{(l)}(t)]dt \\
&\leqq \int_0^x G_j[t, \mathbf{y}(t), \mathbf{y}^{(k)}(t)]dt \sum_{l=k+1}^{\infty} M_j/_{(l-1)!}[\int_0^b K_j(t)dt]^{l-1} \\
&\leqq M_j\{[\int_0^b K_j(t)dt]^k/_{k!} \int_0^b G_j[t, \mathbf{y}(t), \mathbf{y}^{(k)}(t)]dt\} \exp[\int_0^b K_j(t)dt].
\end{aligned}$$

By Q, 2. 3 the term in brackets is the k-th term of a convergent series. Therefore

2. 14 $$\operatorname*{Limit}_{k\to\infty} \int_0^x f_j[t, \mathbf{y}^{(k)}(t)]dt = \int_0^x f_j[t, \mathbf{y}(t)]dt.$$

Hence

$$y_j(x) = \int_0^x f_j[t, \mathbf{y}(t)]dt + a_j, \qquad (j = 1, 2, \cdots).$$

$$\mathbf{y}(x) = \int_0^x \mathbf{f}[t, \mathbf{y}(t)]dt + \mathbf{a} \quad \text{and a solution exists.}$$

To prove uniqueness assume there exists another solution, $\mathbf{y}'(x)$, on $\mathbf{D}^*(x)$. Then

$$\mathbf{y}(x) - \mathbf{y}'(x) \equiv \int_0^x \{\mathbf{f}[t, \mathbf{y}(t)] - \mathbf{f}[t, \mathbf{y}'(t)]\}dt,$$

$$H_j[\mathbf{y}(x) - \mathbf{y}'(x)] \leqq \int_0^x K_j(t) H_j[\mathbf{y}(t) - \mathbf{y}'(t)]dt \leqq M_j \int_0^x K_j(t)\,dt.$$

It follows by induction that

$$H_j[\mathbf{y}(x) - \mathbf{y}'(x)] \leqq M_j/_{n!}[\int_0^b K_j(t)\,dt]^n$$

where n is any positive integer. Hence $H_j[\mathbf{y}(x) - \mathbf{y}'(x)] \equiv 0$ and

$$|y_j(x) - y_j'(x)| \leqq \int_0^x G_j[t, \mathbf{y}(t), \mathbf{y}'(t)] H_j[\mathbf{y}(t) - \mathbf{y}'(t)]dt \equiv 0.$$

Therefore $\mathbf{y}'(x) = \mathbf{y}(x)$ and solution is unique.

The proof of the existence theorem goes thru, with slight changes in the method of proof, if we require that the hypotheses Q, II, III, IV, V are necessarily true only when the vector arguments $\mathbf{w}(x)$, $\mathbf{w}'(x)$, $\mathbf{w}^{(l)}(x)$,

$\boldsymbol{w}'^{(l)}(x)$ are on $\boldsymbol{D}^*(x)$ and are absolutely continuous on $(0, b)$ and add the hypothesis that the function defined by the series Q, 2.3 is absolutely continuous on $(0, b)$. Sufficient conditions that Q, 2.4 and Q, 2.5 follow from the other hypotheses of the set Q are

2.15 $$|f_j(x, \boldsymbol{w}) - f_j(x, \boldsymbol{w}')| \leqq G_j(x, \boldsymbol{w}, \boldsymbol{w}')H_j(\boldsymbol{w} - \boldsymbol{w}'),$$

2.16 $$H_j[\boldsymbol{f}(x, \boldsymbol{w}) - \boldsymbol{f}(x, \boldsymbol{w}')] \leqq K_j(x)H_j(\boldsymbol{w} - \boldsymbol{w}').$$

THEOREM VI. *Given the set of hypotheses Q and in addition* $\boldsymbol{D}^*(x) = \boldsymbol{R}$, *there exists a unique solution of the system*

$$\boldsymbol{y}_n(i) = \sum_{r=0}^{i-1} \boldsymbol{B}_n[r, \boldsymbol{y}_n(r)] + a, \quad y_n(0) = a.$$

That solution exists is a consequence of Q, I and Q, II. $y_n(0)$ *determines* $y_n(1)$ *which by Q, II is in region* $\boldsymbol{R}$. $y_n(0)$, $y_n(1)$ *determine* $y_n(2)$, *also in* $\boldsymbol{R}$. *It follows then by induction that* $y_n(i)$ *is in R and is unique.*

THEOREM VII. *Given*

1. *The set of hypotheses Q,*
2. $G_j(x, \boldsymbol{w}, \boldsymbol{w}') \leqq G_j(x)$, $G_j(x)$ *a positive valued summable function on* $(0, b)$,
3. $H_j[\boldsymbol{f}(x, \boldsymbol{a})]$ *summable on* $(0, b)$, $(j = 1, \cdots)$,
4. *A method of fine subdivision of the interval* $(0, b)$.

Conclusion: $\lim_{n\to\infty} \boldsymbol{y}_n(x) = \boldsymbol{y}(x)$ uniformly in x on $(0, b)$, $[\boldsymbol{y}(x), \boldsymbol{y}_n(x)$ solutions of the differential and difference systems respectively].

Proof. Consider a fixed value of x and for each value of n the subdivision containing x, i. e., $t_{in} \leqq x < t_{i+1n}$.

2.17 $$y_j(x) - y_{nj}(x) = \int_0^x \{f_j[t, \boldsymbol{y}(t)] - f_j[t, \boldsymbol{y}_n(t)]\}dt + \int_{t_{in}}^x f_j[t, \boldsymbol{y}_n(t)]dt$$

By hypothesis Q, 2.4, if $\boldsymbol{w}(x)$ and $\boldsymbol{w}'(x)$ are any pair of vectors on R,

$$\left| \int_{x_1}^{x_2} \{f_j[t, \boldsymbol{w}(t)] - f_j[t, \boldsymbol{w}'(t)]\}dt \right| \leqq M_j \int_{x_1}^{x_2} G_j(t)\,dt.$$

Taking $\boldsymbol{w}'(x) = \boldsymbol{a}$

$$\left| \int_{x_1}^{x_2} f_j[t, \boldsymbol{w}(t)]dt \right| \leqq \int_{x_1}^{x_2} [M_j G_j(t) + |f_j(t, \boldsymbol{a})|]dt.$$

Hence given any positive number $\epsilon > 0$, there exists an integer $n_{\epsilon j} > 0$ such

that for all values of $n > n_{\epsilon j}$, $\left| \int_x^{t_{\epsilon n}} f_j[t, \mathbf{y}_n(t)]dt \right| < \epsilon$. Therefore

2.18 $$| y_j(x) - y_{nj}(x) | \leqq \int_0^x G_j(t) H_j[\mathbf{y}(t) - \mathbf{y}_n(t)]dt + \epsilon, \quad n > n_{\epsilon j}.$$

Also $\mathbf{y}(x) - \mathbf{y}_n(x) = \int_0^x \{\mathbf{f}[t, \mathbf{y}(t)] - \mathbf{f}[t, \mathbf{y}_n(t)]\}dt + \int_{t_{\epsilon n}}^x \mathbf{f}[t, \mathbf{y}_n(t)]dt.$

2.19 $$H_j[\mathbf{y}(x) - \mathbf{y}_n(x)] \leqq \int_0^x K_j(t) H_j[\mathbf{y}(t) - \mathbf{y}_n(t)]dt + H_j\{ \int_x^{t_{\epsilon n}} \mathbf{f}[t, \mathbf{y}_n(t)]dt\}.$$

By hypothesis Q, line 2.5 if $\mathbf{w}(x)$ and $\mathbf{w}'(x)$ are any pair of vectors on $\mathbf{D}^*(x)$

$$H_j\{ \int_{x_1}^{x_2} \mathbf{f}[t, \mathbf{w}(t)]dt - \int_{x_1}^{x_2} \mathbf{f}[t, \mathbf{w}'(t)]dt\} \leqq M_j \int_{x_1}^{x_2} K_j(t)dt.$$

Take $\mathbf{w}'(x) = \mathbf{a}$. Then by 1.8 and 1.9

$$H_j\{ \int_{x_1}^{x_2} f[t, \mathbf{w}(t)]dt\} \leqq M_j \int_{x_1}^{x_2} K_j(t)dt + H[\int_{x_1}^{x_2} \mathbf{f}(t, \mathbf{a})dt]$$

$$H_j\{ \int_{x_1}^{x_2} [t, \mathbf{w}(t)]dt\} \leqq M_j \int_{x_1}^{x_2} K_j(t)dt + \int_{x_1}^{x_2} H_j[\mathbf{f}(t, \mathbf{a})]dt.$$

$\operatorname{Limit}_{x_1 \to x_2} H_j\{ \int_{x_1}^{x_2} [t, \mathbf{w}(t)]dt\} = 0$ uniformly in x and on the region R. Therefore given any positive number ϵ there exists a positive integer $n'_{\epsilon j}$ such that

2.20 $$H_j\{ \int_{x_1}^{x_2} \mathbf{f}[t, \mathbf{y}_n(t)]dt\} < \epsilon \quad \text{when } n > n'_{\epsilon j}.$$

Take $n > n_{\epsilon j} + n'_{\epsilon j}$. Then

$$| y_j(x) - y_{nj}(x) | \leqq \int_0^x G_j(t) H_j[\mathbf{y}(t) - \mathbf{y}_n(t)]dt + \epsilon,$$

$$H_j[\mathbf{y}(x) - \mathbf{y}_n(x)] \leqq \int_0^x K_j(t) H_j[\mathbf{y}(t) - \mathbf{y}_n(t)]dt + \epsilon,$$

$$H_j[\mathbf{y}(x) - \mathbf{y}_n(x)] \leqq M_j \int_0^x K_j(t)dt + \epsilon,$$

$$H_j[\mathbf{y}(x) - \mathbf{y}_n(x)] \leqq M_{j/2!} [\int_0^x K_j(t)dt]^2 + [\int_0^x K_j(t)dt + 1]\epsilon.$$

By simple induction

2.21 $$H_j[\mathbf{y}(x) - \mathbf{y}_n(x)] \leqq \epsilon \exp [\int_0^b K_j(t)dt].$$

Then

2.22 $$| y_j(x) - y_{nj}(x) | \leqq \epsilon \{1 + \exp [\int_0^b K_j(t)dt] \int_0^b G_j(t)dt\}, \quad n > n_{\epsilon j} + n'_{\epsilon j}$$

and $\operatorname{Limit}_{n \to \infty} y_{nj}(x) = y_j(x)$ uniformly in x on $(0, b)$. Therefore $\operatorname{Limit}_{n \to \infty} \mathbf{y}_n(x) = \mathbf{y}(x)$.

Theorems V, VI, and VII are extensions of results obtained by H. J. Ettlinger * for linear difference and differential systems.

UNIVERSITY OF TEXAS,
AUSTIN, TEXAS.

* For abstract of paper see *Bulletin of the American Mathematical Society*, Vol. 34 (1928), p. 708.

Generalized Indecomposable Continua.

By P. M. SWINGLE.

1. *Introduction.* In this paper it is proposed to generalize some of the more useful definitions of indecomposable continua and to prove a number of theorems concerning the sets thus defined. The following definitions of indecomposable continua are known to be equivalent:

DEFINITION A. A continuum is *indecomposable* * when it is not the sum of two proper subcontinua (i. e. not the sum of two subcontinua different from itself).

DEFINITION B. A continuum is *indecomposable* † when each proper subcontinuum is a continuum of condensation.‡

DEFINITION C. A continuum is *indecomposable* § when it contains three points between any two of which it is irreducible.¶

In the generalization of the first two of these definitions the following definitions are useful:

DEFINITION. The set W is a *finished sum* of a set of subcontinua (M) of W if the sum of the subcontinua of (M) contain W but no M of (M) is contained in the sum of the remaining subcontinua of (M).

DEFINITION. A set $W_1, W_2, \cdots, W_k$ of subcontinua of a continuum W will be called a *convergable sequence of* W *under* k when $W_1 = W$, $W_{i-1} \neq W_i$ $(i = 2, \cdots, k)$ although W_{i-1} contains W_i, and $(W_{i-1} - W_i)' \neq W_{i-1}$ but for every subcontinuum C of W_k $(W_k - C)' = W_k$.

Definitions A and B can now be generalized as follows:

DEFINITION A. *A continuum* W *is indecomposable under index* k, where k is any positive integer, when W is the finished sum of k subcontinua of W but is not the finished sum of $k+1$ such subcontinua.

---

* Z. Janiszewski and C. Kuratowski, "Sur les continus indécomposables," *Fundamenta Mathematicae*, Vol. 1, p. 210.

† Z. Janiszewski and C. Kuratowski, *loc. cit.*, theorem II, p. 212.

‡ If M is a point set, M' will be used to denote the set composed of M and the limit points of M. A subcontinuum K of a continuum C is said to be a *continuum of condensation of* C when $(C - K)' = C$.

§ Z. Janiszewski and C. Kuratowski, *loc. cit.*, theorem IV (III), p. 215.

¶ A continuum C is *irreducible between two points* a and b of C if no proper subcontinuum of C contains $a + b$.

DEFINITION B. A continuum W is *convergable under index k,* where k is any positive integer, when W contains a sequence convergable under k but does not contain a sequence convergable under a higher finite index.

It is seen that an indecomposable continuum is both indecomposable and convergable under index one. And it is readily seen that there exist examples which are both indecomposable and convergable under a higher index.

In this paper the impossibility of generalizing definition C in a certain manner will be shown. As it is based upon the idea of continua irreducible between two points an obvious possible generalization would be based upon the following definition:

DEFINITION. If N is a finite subset of a continuum M, then M will be said to be a *continuum irreducible between the points of N* when no proper subcontinuum of M contains N but there exist a proper subcontinuum of M containing any proper subset of N.*

While the proofs of the following theorems are given for the plane similar proofs can be given for a Euclidean space of n dimensions.

2. *Continua Indecomposable under Index k.*

LEMMA 1. *If Z is a connected subset of the connected set C such that $C - Z = C_1 + C_2 + \cdots + C_k$ separate,† then $C_i + Z$ $(i = 1, 2, \cdots, k)$ is connected.*

Assume that $C_i + Z = X + Y$ separate. As Z is connected either X or Y contains it. Consider for example the case where X does. Then $C = (C_1 + C_2 + \cdots + C_{i-1} + C_{i+1} + \cdots + C_k + X) + Y$ separate which is a contradiction. Thus $C_i + Z$ must be connected.

LEMMA 2. *If C is a circle, I is the interior of C, and M is a continuum then $C + M \times I$ is also a continuum; also every point of $M \times I$ is contained in a connected subset of $M \times I$ which has a limit point in C.*

It is evident that $C + M \times I = Z$ is closed. Assume that $Z = X + Y$ separate. Let $U = M - M \times I'$ and let $X \times M = x$ and $Y \times M = y$. As C is connected either X or Y contains it. Consider for example the case where X does. Then x contains $C \times M$ and $y \times C = 0$. Thus $M = (U + x) + y$ separate which is a contradiction. Therefore $C + M \times I$ is a continuum. And since it is a bounded continuum every point of $M \times I$ must be contained in a connected subset of $M \times I$ which has a limit point in C.

* See W. A. Wilson, "On the Separation of the Plane by Irreducible Continua," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), p. 734.

† By $W = W_1 + W_2 + \cdots + W_k$ *separate* is meant that W is the sum of the distinct non-vacuous sets W_i no one of which contains a limit point of any other.

THEOREM 1. *If the continuum W, indecomposable under index k, is the finished sum of the set of subcontinua v_i $(i = 1, 2, \cdots, k)$ then $[v_i - (v_1 + v_2 + \cdots + v_{i-1} + v_{i+1} + \cdots + v_k) \times v_i]'$ is a subcontinuum of W.*

Let T_j $(j = 1, 2, \cdots, q; q < k)$ be the set of maximal connected subsets of $v_1 + v_2 + \cdots + v_{i-1} + v_{i+1} + \cdots + v_k = V_i$. Assume that $W - T_1 = t_1 + \cdots + t_p$ separate. Then it is necessary that p be less than $k + 1$ for if not it is possible to take $p = k + 1$ and then W is the finished sum of the $k + 1$ subcontinua $T_1 + t_g$ $(g = 1, 2, \cdots, p)$, by lemma 1, which is contrary to the definition of a continuum indecomposable under index k. Hence $W - T_1$ can be taken as $t_1 + t_2 + \cdots + t_g$ separate, where every t_i is a connected set. As $W - T_1$ contains T_2 consider for example the case where t_1 contains T_2. Every point of $t_1' - t_1$ is a limit point of $v_i - V_i \times v_i$ as no point of T_1 is a limit point of any other T_j. Then just as above, $t_1' - T_2$ is the sum of a finite number, less than k, of maximal connected subsets which are closed except for limit points in T_2 and which each contain points of t_1 since these sets are separate. Proceeding in this manner it is seen that $(v_i - v_i \times V_i)'$ must be the sum of a finite number, less than k, of maximal connected subsets which are closed except for limit points in V_i. Thus $(v_i - V_i \times v_i)'$ must be a subcontinuum of W otherwise W is not indecomposable under index k.

COROLLARY 1. *If a continuum W is indecomposable under index k then W is the finished sum of a set of subcontinua u_i $(i = 1, 2, \cdots, k)$ of W where every point of u_i which is contained in $u_1 + u_2 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_k = U_i$ is a limit point of $u_i - u_i \times U_i$.*

Since W is indecomposable under index k it is the sum of k subcontinua v_i. And by theorem 1 the points contained in only v_j of the v_i's plus the limit points of this point set is a continuum; that is $u_j = [v_j - v_j \times (v_1 + v_2 + \cdots + v_{j-1} + v_{j+1} + \cdots + v_k)]'$ is a subcontinuum of W. It is thus seen that W is the finished sum of the set of subcontinua u_i where every point of u_i which is contained in U_i is a limit point of $u_i - U_i \times u_i$.

THEOREM 2. *If W is indecomposable under index k then W is the finished sum of a set of k indecomposable subcontinua.*

As W is indecomposable under index k it follows by corollary 1 that W is the finished sum of a set of k subcontinua u_i $(i = 1, \cdots, k)$ where $u_i = [W - (u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_k)]'$. Thus if u_i is decomposable it follows that W is the finished sum of $k + 1$ subcontinua which is a contradiction. Therefore u_i is indecomposable.

LEMMA 3. *If the continuum M, irreducible between two points a and b, is the finished sum of k indecomposable subcontinua u_i $(i = 1, \cdots, k)$ then $u_i = [u_i - u_i(u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_k)]'$.*

Since M is not the sum of $k - 1$ of the u_i's, $(M - u_i)' \neq M$. Thus $(M - u_i)' = A + B$ * where A and B are either proper subcontinua of M or vacuous sets. Let C consist of the points of all those u_j's, except u_i, which have points common with A and let D similarly consist of those having points common with B. As $C + D$ does not contain u_i $C + D \neq M$. Hence $C \times D = 0$. But $(M - C)'$ is a continuum irreducible between two points † as is also $[(M - C)' - D]' = H$. But $H = [M - (C + D)]'$. Thus H is a subcontinuum of u_i but is not a proper subcontinuum of it for then $C + D$ contains u_i and so M is the sum of $k - 1$ of the u_j's which is incorrect. Therefore $[M - (C + D)]' = [u_i - u_i \times (u_1 + \cdots u_{i-1} + u_{i+1} + \cdots + u_k)]' = u_i$.

THEOREM 3. *If the continuum M, irreducible between two points a and b, is the finished sum of k indecomposable subcontinua u_i $(i = 1, \cdots, k)$ then M is indecomposable under index k.*

Assume that M is not indecomposable under index k. Then M is the finished sum of a set of subcontinua v_j $(j = 1, 2, \cdots, k + 1)$. Let v be any one of the v_i's. The set v must contain points contained in only one of the u_i's for every point contained in more than one of the u_i's is a limit point of points contained in only one of the u_i's as shown in lemma 3. And if all of the points of v which are contained in only one u_i are contained in other v_i's then k of the v_i's contain M which is a contradiction under our assumption. Thus v contains a point q which is contained in only one of the u_i's, u say, and which is not contained in any other v_i. As M is irreducible between a and b $(M - u)'$ must be the sum of at most two maximal subcontinua.‡ Consider for example the case where it is the sum of two such subcontinua A and B. If $v \times (A + B) = 0$ then v is a proper subcontinuum of u and so $(u - v)' = u$ and so $(M - v)' = M$ and thus contains q which is a contradiction. Therefore either A or B contains a point of v. Consider for example the case where $A \times v \neq 0$. The set M contains an irreducible continuum T joining A and B.§ But u contains T and so $u = T$. Hence

* C. Kuratowski, "Théorie des continus irréductibles entre deux points," *Fundamenta Mathematicae*, Vol. 3, p. 202, theorem II.

† C. Kuratowski, *loc. cit.*, p. 204, theorem IV.

‡ C. Kuratowski, *loc. cit.*, theorem II, p. 202.

§ C. Kuratowski, *loc. cit.*, theorem VI, p. 205.

if $v \times B \neq 0$, v contains $T = u$ and so there would have to be as many u_i's as v_i's contrary to our assumption. Consider then the case where $v \times B = 0$. If then $(v - v \times A)'$ is a continuum it is a proper subcontinuum of u and so q is contained in some other v_i besides v which is a contradiction. Hence $(v - v \times A)' = X = X_1 + X_2$ separate. And as $A + v$ is a continuum $[M - (A + v)]' = Y$ is a continuum.* And $Y + A + X = M = Y + A + X_1 + X_2$. Also $A + X_j$ $(j = 1, 2)$ is a continuum as $(A + v - A)' = (v - v \times A)'$. Therefore either $A + X_1$ or $A + X_2$ contains a point of Y. Consider for example the case where $A + X_1$ does. The set Y does not contain X_j for if so $Y \times A \neq 0$ although $Y \times q = 0$. Hence $A + X_1 + Y$ is a proper subcontinuum of M containing $a + b$ which is impossible. Therefore M must be indecomposable under index k.

COROLLARY 2. *If M is irreducible between two points then a necessary and sufficient condition that M be indecomposable under index k is that M be the finished sum of a set of k indecomposable subcontinua of M.*

That the condition is necessary follows from theorem 2. And that it is sufficient follows from theorem 3.

THEOREM 4. *If the continuum M, irreducible between two points a and b, is indecomposable under index k, and N is a subset of M irreducibly connected between a and b, then N and $M - N$ are both dense in M.*†

As N' is a continuum irreducible between a and b, $N' = M$ and so N is dense in M.

Assume that $M - N$ is not dense in M. Since M is indecomposable under index k by theorem 2 M is the finished sum of a set of k indecomposable subcontinua u_i $(i = 1, \cdots, k)$. And by lemma 3 every point of M is a limit point of points contained in one of these u_i's alone. Hence since $(M - N)' \neq M$, $(M - N)'$ does not contain every point of M which is contained in only one of the u_i's. Let q then be a point contained only in $u_j = u$ and which is not contained in $(M - N)'$. Let R be a region containing q but such that R' does not contain a point of any u_i besides u nor does it contain a point of $(M - N)'$. Then by lemma 2 $M \times R$ contains a maximal connected subset T which contains q. Then both N and u contain T'. Thus as T' is a proper subcontinuum of u $(u - T')' = u$.

Since N is irreducibly connected between a and b, $N - T = N_1 + N_2$

* C. Kuratowski, *loc. cit.*, theorem III, p. 203.

† See P. Urysohn, "Mémoire sur les multiplicités Cantoriennes," *Fundamenta Mathematicae*, Vol. 8, p. 226, theorem 2.

$+\cdots+N_t$ separate where t is not less than two. Consider for example the case where N_1 contains a and N_2 contains b. Then by lemma 1 $N_1+T+N_2=N$ and hence $t=2$. Thus N_1 and N_2 are each connected.

As every point of T is a limit point of $u-T'$ and so of $N-T'$ every point of T is a limit point of either N_1 or N_2. Hence there must exist a point x which is a limit point of both. Thus $N_1+x+N_2=N$ which is a contradiction. Therefore $M-N$ is dense in M.

3. *Continua Convergable under Index k.* It is evident that if $W_1, W_2, \cdots, W_k$ is a convergable sequence of W under k where W is a continuum convergable under index k that W_j is convergable under index $k-(j-1)$.

THEOREM 5. *If W is convergable under index two then W is irreducible between two points and is the finished sum of two unbounded indecomposable subcontinua.*

Since W is convergable under index two there exists a convergable sequence $W_1=W$, W_2, of W under $k=2$. Hence $(W-W_2)'\neq W$ but every subcontinuum C of W_2 is such that $(W_2-C)'=W_2$. Therefore W_2 is indecomposable.

Let K be a proper subcontinuum of W such that W_2 is a proper subcontinuum of K if such a subcontinuum as K exists. Then $(W-W_2)'$ contains $(W-K)'$ and so $(W-K)'\neq W$ since $(W-W_2)'\neq W$. Thus every subcontinuum k of K is such that $(K-k)'=K$ and so $(K-W_2)'=K$. But $(W-W_2)'$ contains $(K-W_2)'$ and so $(W-W_2)'=W$ which is a contradiction. Therefore there does not exist a subcontinuum of W such as K.

Assume now that $W-W_2=X+Y$ separate. Then W_2+X is a proper subcontinuum such as K. Hence $W-W_2$ must be connected and so $(W-W_2)'$ is a proper subcontinuum of W. Assume that $(W-W_2)'=X+Y$ where X and Y are each proper subcontinua of $(W-W_2)'$. Then neither X nor Y contains $W-W_2$. Either X or Y must contain points of W_2. Consider for example the case where X does. Therefore $X+W_2$ is a proper subcontinuum such as K. As this is a contradiction $(W-W_2)'$ must be indecomposable. If $(W-W_2)'$ is bounded then it contains a strongly connected proper subset* T containing a point of W_2 such that W_2+T contains a subcontinuum such as K. Hence $(W-W_2)'$ must be an unbounded indecomposable proper subcontinuum of W.

* Z. Janiszewski and C. Kuratowski, "Sur les continus indécomposables," *Fundamenta Mathematicae*, Vol. 1, pp. 215-221.

As W_2 contains $[W-(W-W_2)']'$, $[W-(W-W_2)']' \neq W$. Therefore by the above reasoning $[W-(W-W_2)']'$ must also be an unbounded indecomposable proper subcontinuum of W. Hence W is the finished sum of the two unbounded indecomposable proper subcontinua $(W-W_2)'$ and $[W-(W-W_2)']'$.

Let a be any point of $(W-W_2)'$ which is not contained in $[W-(W-W_2)']'$ and let b be any point of $[W-(W-W_2)']'$ which is not contained in $(W-W_2)'$. Then W must be irreducible between a and b otherwise W contains a proper subcontinuum such as K.

As W_2 contains $[W-(W-W_2)']'$, $W_2 = [W-(W-W_2)']'$ otherwise W_2 is a subcontinuum such as K. The following corollary then has been proven:

COROLLARY 3. *If W is convergable under index two and W, W_2 is a convergable sequence of W under two then W is the finished sum of the two unbounded indecomposable proper subcontinua $(W-W_2)'$ and $W_2 = [W-(W-W_2)']'$.*

THEOREM 6. *If W is convergable under index k, $k \neq 1$, then W is the finished sum of k unbounded indecomposable subcontinua.*

Since W is convergable under index k W contains a sequence convergable under k, $W_1 = W, W_2, \cdots, W_k$. Let $W_i = M_{k-(i-1)}$ $(i = 1, 2, \cdots, k)$. Then $W = M_k = (M_k - M_{k-1})' + (M_{k-1} - M_{k-2})' + \cdots + (M_3 - M_2)' + (M_2 - M_1)' + M_1$. Since M_2 is convergable under index two it is known by theorem 5 that $(M_2 - M_1)'$ and M_1 are each unbounded indecomposable proper subcontinua of W. Let j be such that $(M_t - M_{t-1})'$ $(t = 2, 3, \cdots, j)$ is an unbounded indecomposable proper subcontinuum of W but this is not so if $t = j+1$. Assume that M_{j+1} contains a proper subcontinuum M which has M_j as a proper subcontinuum. Hence $(M_{j+1} - M_j)'$ contains $(M_{j+1} - M)'$ and $(M_{j+1} - M_j)'$ also contains $(M - M_j)'$. Thus as $(M_{j+1} - M_j)' \neq M_{j+1}$ $(M_{j+1} - M)' \neq M_{j+1}$ and $(M - M_j)' \neq M$. Thus M_{j+1} which is convergable under index $j+1$ contains the sequence, $M_{j+1}, M, M_j, M_{j-1}, \cdots, M_2, M_1$, convergable under $j+2$. As this is a contradiction a subcontinuum such as M does not exist. It is then necessary, just as shown in the proof of theorem 5, that $(M_{j+1} - M_j)'$ be a continuum but not either a decomposable or a bounded indecomposable continuum. Thus it must be an unbounded indecomposable proper subcontinuum of W. Hence $j = k$ and so W is the finished sum of a set of k unbounded indecomposable proper subcontinua.

THEOREM 7. *Let W be convergable under index k and let it also be the*

finished sum of a set of subcontinua M_i $(i=1,\cdots,k)$. Let K be a subcontinuum of W which contains a point contained in M_b alone and also a point not contained in M_b alone. Then the points of K which are contained in M_b alone are contained in a finite number of subcontinua of $K \times M_b$.

Let Z be composed of all the points of W which are not contained in M_b alone. Let Q be the maximal subcontinuum of $Z+K$ which contains K. As Z contains all M_i's except M_b, $Z \times Q$ contains at most $k-1$ maximal subcontinua. Let $C_1,\cdots,C_q$ be these maximal subcontinua. Assume that $Q-C_1=Q_1+Q_2+\cdots+Q_n$ separate where n is taken greater than k if possible. Let $W_1=W$, $W_2=Q=C_1+Q_1+\cdots+Q_n$, and $W_{n-t+2}=C_1+Q_1+\cdots+Q_t$ $(t=2,\cdots,n)$. If $Q\neq W$ then $(W-Q)'\neq W$ for W is a finished sum of the M_i's. And if $Q=W$, $(W-W_3)'\neq W$ for it equals Q_n'. Also $(W_i-W_{i+1})'\neq W_i$ since $(W_i-W_{i+1})'=Q_{n-i+2}'$. Hence since W is convergable under index k it is seen that $n=k$ at most. Thus $Q-C_1$ must be the sum of a finite number of maximal connected subsets. Let T be one of these maximal connected subsets such that T contains a C_i, C_2 say. Treating $T'-C_2$ in a similar manner it is seen that it must be the sum of a finite number of maximal connected subsets. Thus proceeding in this manner it is seen that $K-Z$ must be the sum of a finite number of maximal connected subsets of M_b from which fact the truth of the theorem follows.

THEOREM 8. *If W is convergable under index k then W is indecomposable under index k.*

By theorem 6 W is the finished sum of k indecomposable subcontinua M_i $(i=1,\cdots,k)$. Assume that W is also the finished sum of the set of subcontinua Z_j $(j=1,\cdots,p)$. Then Z_j contains a point q which is contained in it alone. But q must be contained in at least one M_i, M_n say. And from the proof of theorem 6 it is seen that every point of W is a limit point of the set of points which are contained in one M_i only. Thus Z_j must contain points which are contained in an M_n and Z_j only. Let x be such a point. Then by theorem 7, x is contained in a subcontinuum of $M_n \times Z_j$. Hence as every proper subcontinuum of an indecomposable continuum is a continuum of condensation it is necessary that Z_j contain M_n. Therefore p cannot be greater than k and so W is indecomposable under index k.

THEOREM 9. *Let W be convergable under index k and let it be the finished sum of the k unbounded indecomposable subcontinua M_i. Then no M_j $(j=1,2,\cdots,i-1,\ i+1,\cdots,k)$ contains a point of k maximal*

strongly connected proper subsets * *of* M_i *which contain also points of* M_i *alone.*

If the theorem is not true then M_j contains a point of k maximal strongly connected proper subsets of M_i and so there exist a subcontinuum N of W such that $(W - N)' \neq W$ although N contains a sequence convergable under k and so W contains a sequence convergable under index $k + 1$. As this is a contradiction the theorem must be true.

THEOREM 10. *If* W *is convergable under index* k *then there exist a set of* k *points about which* W *is irreducible.*†

Let M_i be a set of k unbounded indecomposable subcontinua of which W is the finished sum. Let a_i be a point of M_i which is not contained in M_j $(j = 1, \cdots, i - 1, i + 1, \cdots, k)$ nor contained in a maximal strongly connected proper subset of M_i which has a point common with any M_j and there exist such points by theorem 9 and because of the fact that M_i contains an uncountable number of maximal strongly connected proper subsets. If W is not irreducible about the set $a_1 + \cdots a_k$ then there exist a proper subcontinuum K of W which contains it. Therefore by theorem 7, since K cannot contain all of some M_i, M_c say, a_c is contained in a maximal strongly connected proper subset of M_c which contains a point of some M_j, $j \neq c$, which is a contradiction. Hence W is irreducible about k points.

COROLLARY 4. *If* W *is convergable under index* k *then there exist* q *points,* q *not greater than* k*, between which* W *is irreducible.*

There exists by theorem 10 a set of k points about which W is irreducible. It is then evident that this set of points must contain a subset between the points of which W is irreducible.

4. *Continua Irreducible Between* k *Points.* As stated in definition C it has been shown that in order that a continuum M be indecomposable it is necessary and sufficient that M contain a set N of $k + 1$ points, where $k = 2$, such that M is irreducible between any combination of k of these points. Here it will be shown that there does not exist a set having this property for k greater than two.

By definition C it is seen that an indecomposable continuum is irreducible between two points. A number of theorems have been proven giving neces-

* See Z. Janiszewski and C. Kuratowski, "Sur les continus indécomposables," *Fundamenta Mathematicae,* Vol. 1, p. 218.

† The continuum W is said to be irreducible about the set N of W if no proper subcontinuum of W contains N.

sary and sufficient conditions that an irreducible continuum be indecomposable. Here the impossibility of generalizing these, in a certain manner, is shown.

In the remainder of this paper let $N = a_1 + a_2 + \cdots + a_k$ and let M be a continuum irreducible between the points of N. If M is also irreducible between the points of $N - a_i + a$ then a will be called an *a_i co-end-point of M.* The set composed of all the a_i co-end-points of M will be called the *a_i co-end-set of M* and will be represented by $(a_i)_{N-a_i}$ or by (a_i).*

It has been shown that if $k = 2$ a necessary and sufficient condition that the irreducible continuum M be indecomposable is that $(a_1) \times (a_2) \neq 0$.† However this is not true for k greater than two as shown by the following theorem.

THEOREM 11. *If k is greater than two then $(a_i) \times (a_j) = 0$.*

Assume that $(a_i) \times (a_j)$ contains a point a. Let X be any subcontinuum of M containing $N - a_i$ and let Y be any subcontinuum containing $N - a_j$ such that $X \neq M \neq Y$. But $X + Y$ contains N and so $X + Y = M$. Hence either X or Y contains a in which case either $X = M$ or $Y = M$ which is a contradiction. Thus $(a_i) \times (a_j) = 0$.

THEOREM 12. *If Y is a proper subcontinuum of M containing $N - a_i$ then $(M - Y)'$ is a continuum irreducible between a_i and any point of $Y \times (M - Y)'$.*

Assume that $M - Y = X + Z$ separate. Either X or Z contains a_i. Say that Z does. Then the continuum $Z + Y$ contains N and so $Z + Y = M$ which is a contradiction. Hence $M - Y$ is connected and so $(M - Y)'$ is a subcontinuum of M. It is then evident that $(M - Y)'$ is irreducible between a_i and any point of $Y \times (M - Y)'$ otherwise M is not irreducible between the points of N.

THEOREM 13. *If M is irreducible about the points of the finite set Z then $Z \times (a_i) \neq 0$.*

Assume that $Z \times (a_i) = 0$. Let Y_i be composed of all those points of M which are contained in proper subcontinua of M containing $N - a_i$. Let z be any point of Z. There exists a proper subcontinuum of M containing $N - a_i + z$ otherwise $(a_i) \times Z$ contains z. Hence Y_i contains Z. And since Z contains a finite number of points, n say, there exist at least n sub-

* See C. Kuratowski, "Théorie des continus irréductibles entre deux points II," *Fundamenta Mathematicae,* Vol. 10, p. 230. See also P. M. Swingle, "End-sets of continua irreducible between two points," *Fundamenta Mathematicae.*

† This follows readily from the work of S. Mazurkiewicz, Z. Janiszewski and C. Kuratowski. See however P. M. Swingle, *loc. cit.*

continua of Y_i whose sum contains $Z+N-a_i$ and whose sum then equals M. Thus $Y_i=M$ which is a contradiction, as $Y_i \times a_i=0$. Therefore $Z\times(a_i)\neq 0$.

THEOREM 14. *If M is irreducible about the finite point set Z then Z contains a subset of k points between the points of which M is irreducible.*

Since M is irreducible about the finite point set Z, the set Z must contain a subset X between the points of which M must be irreducible. Let x be the number of points in X. By theorem 13 $X\times(a_i)\neq 0$. And by theorem 11 $(a_i)\times(a_j)=0$. Hence X contains at least k points and similarly N must contain at least x points. Hence $x=k$.

COROLLARY 5. *If Y is a proper subcontinuum of M containing $N-a_i$ then, if k is greater than two, $(M-Y)'\neq M$.*

For if $(M-Y)'=M$ then $(M-Y)'$ is irreducible between k points and by theorem 12 it is also irreducible between two points. Hence by theorem 14 $k=2$.

COROLLARY 6. *If k is greater than two then $(a_i)'\neq M$.**

Let Y be a proper subcontinuum of M containing $N-a_i$. Then $Y\times(a_i)=0$. Hence $(M-Y)$ contains (a_i) and so $(M-Y)'$ contains $(a_i)'$. Thus by corollary 5 $(a_i)'\neq M$.

THEOREM 15. *If k is greater than two and Y is a proper subcontinuum of M containing $N-a_i$ then Y contains (a_j) $(j\neq i)$.*

Assume that Y does not contain a of (a_j). Then $(M-Y)'$ contains a. Let K be a proper subcontinuum of M containing $N-a_j$. As k is greater than two $(N-a_i)\times(N-a_j)\neq 0$ and so $K\times Y\neq 0$. Hence $K+Y$ is a subcontinuum of M containing N and so $K+Y=M$. Thus K contains $(M-Y)'$ and so contains a. Hence K contains $a+N-a_j$ and so $K=M$ which is a contradiction. Therefore Y contains a and so contains (a_j).

It has been shown that, for $k=2$, a necessary and sufficient condition that the irreducible continuum M be indecomposable is that $(a_1)+(a_2)=M$.† The following theorem shows the impossibility of generalizing this for the cases where k is greater than two.

THEOREM 16. *If k is greater than two $(a_1)+(a_2)+\cdots+(a_k)\neq M$.*

* For $k=2$ C. Kuratowski has proven that $(a_i)'=M$ is a necessary and sufficient condition that the irreducible continuum M be indecomposable. See C. Kuratowski, *loc. cit.*, *Fundamenta Mathematicae*, Vol. 10, p. 235, theorem I.

† P. M. Swingle, *loc. cit.*

Assume that $(a_1)+\cdots+(a_k)=M$. Let Y be a proper subcontinuum of M containing $N-a_i$. Then $Y\times(a_i)=0$. By theorem 15 Y contains the sum of the (a_j)'s $(j\neq i)$. Hence $Y=(a_1)+\cdots+(a_{i-1})+(a_{i+1})+\cdots+(a_k)$. By theorem 11 $(a_i)\times(a_j)=0$. Hence $M-Y=(a_i)$. Thus (a_i) $(i=1,\cdots,k)$ is not closed while the sum of $k-1$ of the (a_j)'s is closed. As this is a contradiction the theorem is true.

It is possible however to generalize the (a_i) co-end-set for $k=2$ in another manner. Let $Z_i=N-a_i$ and let $(Z_i)_{a_i}=(Z_i)$ be the set of point sets which have the property that together with a_i M is irreducible between these points. The set (Z_i) will be called the Z_i *co-end-set of* M.

It is evident that for $k\neq 2$ $(Z_i)\times(Z_j)\neq 0$ *for every* M.

THEOREM 17. *If Y is a proper subcontinuum of M containing the set Z of (Z_i) then Y contains (Z_i), if k is greater than two.*

Let z be a set of (Z_i) and b a point of Z. Then by theorem 13 $z\times(b)\neq 0$. And by theorem 15 Y contains (b). Hence as the (b)'s contain z, Y contains (Z_i).

THEOREM 18. *If k is greater than two,* $(Z_1)+(Z_2)+\cdots+(Z_k)\neq M$.

Let Z be a set of (Z_i). Then M is irreducible between the points of $Z+a_i$. Then $(a_1)+\cdots+(a_k)$ contains $Z+a_i$ by theorems 13 and 14. Hence $(a_1)+\cdots+(a_k)\neq M$ contains $(Z_1)+\cdots+(Z_k)$ and so $(Z_1)+\cdots+(Z_k)\neq M$.

THEOREM 19. *If B is a set of $k+1$ points of a continuum C, where $k\neq 1$ or 2, then C is not irreducible between every combination of k of the points of B.*

Assume that C is irreducible between every combination of k of the points of B. Then C is not irreducible between the combination of any $k-1$ of the points of B for assume that Z is a set of $k-1$ of these points. Then by theorem 14 Z must contain k points. Thus each combination of $k-1$ of the points of B is contained in a proper subcontinuum K of M. Consider the sum S of a set of such K's where S contains and only contains a K for each combination of the points of $B-b_{k+1}$, where $B=b_1+\cdots+b_{k+1}$, taken $k-1$ at a time and where b_1 is contained in each of these combinations. Then S is a subcontinuum of C containing k points of B. Hence $S=M$. Thus one of the K's, K_1 say, contains b_{k+1} and so contains k of the points of B. Therefore $K_1=M$ which is a contradiction. Hence C is not irreducible between every combination of the points of B taken k at a time.

OHIO STATE UNIVERSITY.

Differential Equations Containing Absolute Values of Derivatives.

By Cletus Odia Oakley.

1. *Introduction.* In this paper we consider the non-linear differential equation

$$(1) \qquad u'' + p_1 u' + p_2 u + q_1 \mid u' \mid + q_2 \mid u \mid = \phi,$$

where p_1, p_2, q_1, q_2, ϕ are real, single-valued and continuous functions of the real variable x throughout an interval (a, b); and investigate the questions of existence and nature of real solutions, separation of zeros, comparison and oscillation theorems.* Although equation (1) is non-linear, it will be referred to as a linear differential equation with absolute values for, as we shall see, the solutions of it follow those of a set of related linear equations, namely:

$$(2) \qquad \begin{array}{ll} (2_1) u'' + p_1 u' + p_2 u + q_1 u' + q_2 u = \phi, & u(+),\dagger\ u'(+); \\ (2_2) u'' + p_1 u' + p_2 u - q_1 u' + q_2 u = \phi, & u(+),\ u'(-); \\ (2_3) u'' + p_1 u' + p_2 u - q_1 u' - q_2 u = \phi, & u(-),\ u'(-); \\ (2_4) u'' + p_1 u' + p_2 u + q_1 u' - q_2 u = \phi, & u(-),\ u'(+). \end{array}$$

These equations will be known as the associated equations or the set associated with (1).

In § 2 we establish the existence of solutions of (1) with continuous first and second derivatives. We also lay down certain definitions, call attention to the first order homogeneous equation and state without derivation two theorems which will be of use in the development of the subject and which are readily established. The case of constant coefficients is treated in § 3, where, in view of the richness of the material, considerable space is devoted to the theory centering around the homogeneous equation. In § 4 we treat the case of variable coefficients where the equation is again homogeneous, deriving identities and theorems which are in some respects similar to those of the Sturmian theory associated with linear equations. While the treatment is such that the classic theory can be obtained by making identically zero the coefficients of the terms involving absolute values, yet it should be pointed out that certain relations (which exist when these coefficients are

* See Bôcher, "Leçons sur les Méthodes de Sturm," where further references to the linear theory will be found.

† By a symbol $\mu(+)$ we shall mean μ is positive-valued.

identically zero) can not exist when the absolute-value terms are actually present and hence that much of the theory herein treated has no counterpart in the classic material. It is in this section that the prnicipal results of the paper occur.

2. *Existence and Nature of Solutions.* We begin with the following

DEFINITION. *A function $u(x)$, continuous in its first two derivatives in the interval (a, b), which satisfies the equation (1) will be said to be a solution of (1) in that interval.*

The interval may be open or closed. A solution of the first order equation would be defined in like manner. It will be instructive to begin the discussion of the existence and nature of solutions by examining the first order equation

$$u' + pu + q\,|\,u\,| = \phi \tag{3}$$

and its two associated equations

$$(4)\qquad \begin{array}{ll} (4_1) & u' + pu + qu = \phi, \\ (4_2) & u' + pu - qu = \phi. \end{array}$$

In the first place, if $\phi \equiv 0$, a solution u_1 of (4_1) such that $u_1(\gamma) = \delta(+)$, where γ is any point in the interval (a, b), will remain positive throughout (a, b) and will, moreover, satisfy (3). In a similar manner a solution u_2 of (4_2) with initial conditions $u_2(\gamma) = \delta(-)$ will remain negative and will be a solution of (3). Now suppose $\phi \not\equiv 0$ and consider u_1 of (4_1) such that $u_1(x_1) = \delta(+)$. To the right of x_1 and so long as it remains positive, this function $u_1(x)$ will satisfy (3). Let $u_1(x_2) = 0$, $\phi(x_2) \neq 0$, x_2 being the first vanishing point of u_1 to the right of x_1. A solution $u_2(x)$ of (4_2)* with the condition $u_2(x_2) = 0$ will match up with u_1 with continuous derivative for at x_2 both (4_1) and (4_2) reduce to $u'(x_2) = \phi$, with $u(x_2) = 0$. Continuing the process of piecing together these two functions $u_1(x)$ and $u_2(x)$, we obtain a solution $u(x)$, $u(\gamma) = \delta$, of (3), throughout (a, b).

We turn now to equation (1) and the set (2) and confine our attention to what will be a typical example of the manner in which a solution of (1) may be obtained by matching up the solutions of (2). Consider, then, a solution u_1 of (2_1) with initial conditions $u_1(x_1) = 0$, $u_1'(x_1) = \delta'(+)$. So long as u_1' remains positive, u_1 will also be a solution of (1). Equation (1), therefore, possesses a solution at least in the open interval (x_1, x_2) where

* The discussion here is typical rather than complete. If $\phi(x_2) = 0$ so that $u_1'(x_2) = 0$, it may happen that $u_1(x)$ of (4_1) continues to be a solution of (3).

x_2 is the first vanishing point of u_1' to the right of x_1. Beyond this point the solution of (2_1) will continue to be a solution of (1) if and only if u_1' remains positive in the neighborhood of x_2. We suppose, then that u_1' is negative to the right of x_2 and consider a solution u_2 of the equation (2_2) such that $u_2(x_2) = u_1(x_2)$, $u_2'(x_2) = 0$. It is clear that u_2 will be a solution of (1) if and only if u_2' is negative in the right neighborhood of x_2. In order to show that u_2' is negative to the right of x_2, we notice first that at x_2 we have, from equation (2_1), $u''(x_2) + p_2u(x_2) + q_2u(x_2) = \phi$ and that $u_1''(x_2)$ is negative. Also, at this point, we get, from equation (2_2), $u''(x_2) + p_2u(x_2) + q_2u(x_2) = \phi$ and the solution u_2 defined above is such that $u_2(x_2) = u_1(x_2)$, $u_2'(x_2) = u_1'(x_2) = 0$. Hence $u_2''(x_2)$ is also negative and the solution $u(x)$ with continuous first and second derivatives of equation (1) persists, through the medium of (2_1) and (2_2), in the open interval (x_1, x_3) where x_3 is the point to the right where first u_2 or u_2' becomes zero. If u_2' is the first to become zero, then the solution $u(x)$ of (1) may switch back to that of (2_1); but if u_2 is the first to become zero, then, if we consider a solution u_3 of (2_3) such that $u_3(x_3) = 0$, $u_3'(x_3) = u_2'(x_3)$, a repetition of the argument carries the solution of (1) on by means of equation (2_3) and further, perhaps,* by equation (2_4). Hence, by means of the solutions of the associated set (2), the existence of a unique solution with continuous first and second derivatives of equation (1)† is demonstrated. Moreover this is in essence a general solution according to the following

DEFINITION. *A solution $u(x)$ of (1) will be called a general solution in (a, b) if $u(x)$ contains arbitrary elements such that for every x_0 in (a, b) and for any whatever given set of constants δ, δ' it is possible to specialize those arbitrary elements so as to make $u(x_0) = \delta$, $u'(x_0) = \delta'$.*

For suppose we take any point γ, fixed in (a, b), and contemplate the set of solutions $v(x)$, $v(\gamma) = \sigma$, $v'(\gamma) = \sigma'$, where σ, σ' are arbitrary real numbers; and further consider any point x_0 in (a, b) and two constants δ, δ'. Now we have just shown that there exists a solution $u(x)$, $u(x_0) = \delta$, $u'(x_0) = \delta'$, which, if followed to the point γ, will be such that $u(\gamma) = \sigma_1$,

* A solution of (1) could, of course, follow those of (2_1), (2_2) and (2_3) only; or of (2_1) and (2_3) only; etc.

† By similar methods the existence of a solution continuous in its first n-derivatives of the equation

$$(1) \qquad u^{(n)} + p_1u^{(n-1)} + \cdots + p_nu + q_1 \mid u^{(n-1)} \mid + q_2 \mid u^{(n-2)} \mid + \cdots + q_n \mid u \mid = \phi$$

could be demonstrated.

$u'(\gamma)=\sigma_1'$. Hence by specializing σ, σ' to σ_1, σ_1', we obtain from $v(x)$ the solution $u(x)$, $u(x_0)=\delta$, $u'(x_0)=\delta'$. Thus $v(x)=f(\gamma;\ \sigma,\ \sigma',\ x)$ is the general solution and the arbitrary elements are σ, σ'. This does not, however, give us the desired information concerning the way in which the arbitrary elements enter into the general solution. We notice that even if u_1 is a solution of the homogeneous equation

$$u''+p_1u'+p_2u+q_1\,|\,u'\,|+q_2\,|\,u\,|=0, \tag{1'}$$

it does not follow that ku_1, where k is a constant,* will be a solution; nor will the sum of two solutions (u_1+u_2) in general be a solution. Although the arbitrary elements evidently do not enter linearly, the notion of linearly dependent solutions is an important one. We have immediately the

THEOREM I. *A necessary and sufficient* † *condition that two solutions of* (1′) *be linearly dependent is that the wronskian vanish identically.*

By methods that are standard in the treatment of linear equations the following theorem can be established.

THEOREM II. *Any solution of* (1′) *which together with its first derivative vanishes at a point* γ, *vanishes identically.*‡

We shall exclude this solution from consideration throughout this paper.

The general, positive-valued solution u_1 of the homogeneous equation

$$u'+pu+q\,|\,u\,|=0 \tag{3'}$$

is

$$u_1=c_1\exp\left[-\int_{c_0}^{x}\{p(\xi)+q(\xi)\}d\xi\right]$$

where c_1 is a positive constant and c_0 is any point in (a,b); and the general, negative-valued solution u_2 is

$$u_2=c_2\exp\left[-\int_{c_0}^{x}\{p(\xi)-q(\xi)\}d\xi\right]$$

where c_2 is a negative constant. Hence

$$u_2=(c_2/c_1)\exp\left[2\int_{c_0}^{x}q(\xi)\,d\xi\right]u_1.$$

---

* In general ku_1 will be a solution of (1′) when and only when k is positive.

† See Bôcher, "Certain Cases in Which the Vanishing of the Wronskian is a Sufficient Condition for Linear Dependence," *Transactions of the American Mathematical Society*, Vol. 2 (1901), p. 139. Theorem I °(p. 140) shows that $u_1=cu_2$ provided $u_2\not\equiv 0$. If $u_2=0$, then $u_1=0$ necessarily and we may use the same constant c. Finally, from the continuity of u_1', u_2' (in particular at the zeros of u_2), it follows that we must have the same c throughout (a,b) even though u_2 changes sign in (a,b).

‡ Theorems corresponding to I and II could be formulated for equation (1), $\phi\equiv 0$, in footnote †, page 661.

The general, complex-valued solution $z(x) \doteq u(x) + iv(x)$, $u(x)$ and $v(x)$ real, and $v(x) \not\equiv 0$, is*

$$z = u + iv = k_1 \exp\left[-\int_{c_0}^{x} \{p(\xi) + q(\xi)\} d\xi\right]$$
$$-(k_2^2/k_1) \exp\left[-\int_{c_0}^{x} \{p(\xi) - q(\xi)\} d\xi\right] + 2ik_2 \exp\left[-\int_{c_0}^{x} p(\xi)\, d\xi\right]$$

where $k_1 > 0$ and k_2 is real and not zero. We call attention to the fact that the first term of z is the general, positive-valued solution, that the second term is the general, negative-valued solution, and that the last term is a solution of the ordinary linear equation. If $q \not\equiv 0$, equation (3′) can not have a pure imaginary solution.

3. *The Case of Constant Coefficients.* We write down the equation

$$u'' + a_1 u' + a_2 u + b_1 |u'| + b_2 |u| = 0 \tag{5}$$

and the associated set

$$(6)\quad \begin{array}{ll} (6_1) & u'' + a_1u' + a_2u + b_1u' + b_2u = 0 \\ (6_2) & u'' + a_1u' + a_2u - b_1u' + b_2u = 0 \\ (6_3) & u'' + a_1u' + a_2u - b_1u' - b_2u = 0 \\ (6_4) & u'' + a_1u' + a_2u + b_1u' - b_2u = 0. \end{array}$$

Along with (5) and (6) we shall need to consider the quantities

$$\begin{aligned} m_{11},\dagger\ m_{12} &= 1/2\, \{-(a_1 + b_1) \pm [(a_1 + b_1)^2 - 4(a_2 + b_2)]^{1/2}\} \\ m_{21}, m_{22} &= 1/2\, \{-(a_1 - b_1) \pm [(a_1 - b_1)^2 - 4(a_2 + b_2)]^{1/2}\} \\ m_{31}, m_{32} &= 1/2\, \{-(a_1 - b_1) \pm [(a_1 - b_1)^2 - 4(a_2 - b_2)]^{1/2}\} \\ m_{41}, m_{42} &= 1/2\, \{-(a_1 + b_1) \pm [(a_1 + b_1)^2 - 4(a_2 - b_2)]^{1/2}\} \\ \Delta_1 &= (a_1 + b_1)^2 - 4(a_2 + b_2) \\ \Delta_2 &= (a_1 - b_1)^2 - 4(a_2 + b_2) \\ \Delta_3 &= (a_1 - b_1)^2 - 4(a_2 - b_2) \\ \Delta_4 &= (a_1 + b_1)^2 - 4(a_2 - b_2) \end{aligned}$$

which are the roots and discriminants respectively of the characteristic equations corresponding to the set (6). If $\Delta_j < 0$, we write $m_{j1} = \alpha_j + i\beta_j$,

---

* If $z = u + iv$ is to satisfy (3′) we must have, (a) $v' + pv = 0$, (b) $u' + pu + q(u^2 + v^2)^{1/2} = 0$ simultaneously. Equation (a) is immediately integrable. Equation (b) represents a particular type of non-linear equation in which, because of the peculiar way in which v enters, the variables are made separable by making the transformation $u = gt$ and afterwards choosing g in such wise as to satisfy the equation $g' + pg = 0$.

† Read m_{j_1} with $+$ sign, m_{j_2} with $-$ sign; $(j = 1, 2, 3, 4)$.

$m_{j2} = \alpha_j - i\beta_j$, $j = 1, 2, 3, 4$; $i = (-1)^{1/2}$. We shall designate by y_j, v_j, s_j, t_j, real solutions of (6_j) as follows:

$$\text{(Exp.)} \qquad s_j \begin{cases} (7) & y_j = c_1 e^{m_{j1}x} + c_2 e^{m_{j2}x}, \quad \Delta_j > 0,\ m_{j1} \neq m_{j2} \\ (8) & v_j = c_1 e^{m_j x} + c_2 x e^{m_j x}, \quad \Delta_j = 0,\ m_{j1} = m_{j2} = m_j \end{cases}$$

$$\text{(Trigonometric)} \qquad (9) \quad t_j = e^{\alpha_j x} c_1 \sin(\beta_j + c_2),\ \Delta_j < 0,$$

where c_1, c_2 are the arbitrary elements and s_j represents either exponential function y_j or v_j.

It is convenient at this point to consider some properties of (7) and (8) that are of fundamental importance to our problem. In the first place, suppose in (7) that m_{j1}, and m_{j2} are of the same sign. Then y_j, with appropriate initial conditions to make c_1 and c_2 of one and the same sign, will represent a solution of (5) throughout any finite interval (a, b). This follows at once because in this case neither y_j nor y_j' is ever zero and this solution of (5) with these boundary conditions remains forever on a solution of one and the same associated equation (6_j). If boundary conditions necessary to make c_i, c_2 of opposite sign are imposed, then y_j vanishes at one and only one point x_0 and y_j' vanishes not at all. In this case y_j will represent a solution of (5) to the right (or left) of x_0. Similarly, if m_{j1} and m_{j2} are of opposite sign, then a solution y_j of (7), with initial conditions such that c_1 and c_2 are of opposite sign, will vanish at no point while its derivative y_j' vanishes at one and but one point x_1. For y_j to vanish at one and only one point x_0 and also for y_j' to vanish at one and only one point x_1 it is both necessary and sufficient that c_1, c_2 be of opposite sign and that at the same time m_{j1}, m_{j2} be of the same sign. Further, if m_{j1}, m_{j2} are negative, then x_0 will lie to the left of x_1; if m_{j1}, m_{j2} are positive, x_0 will lie to the right of x_1. Moreover, the distance $(x_1 - x_0) = (\log m_{j1} - \log m_{j2})/(m_{j2} - m_{j1})$, is not zero and is independent of c_1, c_2. Examining v_j and assuming $c_2 m_j \neq 0$, we note: that there always exist two and only two points x_0 and x_1 such that $v_j(x_0) = 0$, $v_j'(x_1) = 0$; that if m_j is negative, x_0 will lie to the left of x_1; that if m_j is positive, x_0 will lie to the right of x_1; that $(x_1 - x_0) = 1/m_j$, is not zero and is independent of c_1, c_2. Hence any function $s_j(x)$ of the form (7) or (8) will possess the property that $s_j(x_0) = 0$, $s_j'(x_1) = 0$, provided that m_{j1}, m_{j2} are of the same sign and c_1, c_2 are of opposite sign.

With these things in mind, we seek conditions on the coefficients of (5) so that a solution $u(x)$, $u(\gamma) = \delta(+)$, $u'(\gamma) = \delta'(+)$ of (5) shall be made up entirely of solutions $s_j(x)$ and which shall have as many zeros as possible in the interval from minus infinity to plus infinity. We shall consider that δ, δ' have been so chosen that c_1, c_2 are of opposite sign. Two things

are evident at once:—First: In order that zeros of u exist to the left of γ, s_1 must vanish to the left of γ, say $s_1(x_0)=0$, where $x_0<\gamma$. Second: In order that zeros of u exist to the right of γ, s_1' must vanish to the right of γ, say $s_1'(x_1)=0$, $x_1>\gamma$. And it follows immediately from the considerations above that m_{11} and m_{12} must both be of the same sign and negative. Assuming that m_{11} and m_{12} are negative, we follow this function s_1 up to the point x_1 and there connect with a solution of (6_2).

The solution $u(x)$ of (5), then, will follow s_1 up to x_1 where it will join to $s_2(x)$ with initial conditions $s_2(x_1)=s_1(x_1)$, $s_2'(x_1)=s_1'(x_1)=0$. If s_2 is to vanish to the right of x_1, say at x_2, it follows that m_{21}, m_{22} must both be positive. We assume this condition on m_{21}, m_{22} and note that this gives $u(x)$ two zeros so far—one at x_0 and one at x_2. In order for s_3 to join with s_2 at x_2 and to be such that $s_3'(x_3)=0$, $x_3>x_2$, it is necessary that m_{31}, m_{32} be negative. Now if s_4 matches up with s_3 at x_3 and vanishes to the right at x_4, say, it is obvious that s_1, s_2, s_3, s_4 will continue to match up throughout any finite interval (a,b) and form an oscillating solution $u(x)$ of (5). But the condition for s_4' to vanish at x_3 and for s_4 to vanish at x_4 is that m_{41}, m_{42} be positive and this is impossible in view of the conditions already imposed on m_{11}, m_{12}, m_{21}, m_{22}, m_{31}, m_{32}. We have therefore established the following

THEOREM III. *There exists no solution $u(x)$ of (5) made up wholly of exponential parts $s_j(x)$ which vanishes more than twice in any interval (a,b).*

Because of the fact that a solution of (5) may possess two and only two zeros we lay down the following

DEFINITION. *A solution $u(x)$ of (5) will be said to oscillate in an interval (a,b) if, in that interval, it vanishes as many as three times.*

While the totality of conditions

$$m_{1\nu}(-),\ m_{2\nu}(+),\ m_{3\nu}(-),\ m_{4\nu}(+), \qquad \nu=1,2 \tag{10}$$

is not realizable, yet the *order of the signs exhibited* is of highest importance. For suppose we follow a solution $u(x)$ as we did (in the above discussion) up to the point x_3 and ask if at this point the solution could be continued by t_4 of trigonometric form (9)? An examination of the m's shows that Δ_4 can not be less than zero when $m_{11}(-)$, $m_{12}(-)$, $m_{21}(+)$, $m_{22}(+)$, $m_{31}(-)$, $m_{32}(-)$. It is not difficult to show that it is impossible for any Δ_j to be negative when the other m's follow the law of signs (10). Hence the

THEOREM IV. *There exists no oscillating solution $u(x)$ of (5) made up of three exponential parts and one trigonometric part.*

Suppose, now, $m_{11}(-)$, $m_{12}(-)$, $m_{21}(+)$, $m_{22}(+)$ and that we have followed through on a solution u up to the point x_2. If $\Delta_3 < 0$, a solution t_3 of (6_3) and of form (9) will connect with s_2 at x_2. It is obvious that t_3' will vanish to the right of x_2 at some point which we shall call x_3. Further, if $\Delta_4 < 0$, t_4 will match up with t_3 at x_3 and will vanish to the right at x_4 from which point, it is apparent, s_1, s_2, t_3, t_4 will continue to match up to form an oscillating solution of (5). This establishes the

THEOREM V. *There exists a class of oscillating solutions $u(x)$ of (5) members of which are made up of two exponential parts (where u is positive-valued) and two trigonometric parts (where u is negative-valued).* It may easily be verified that: *There exists another class of oscillating solutions $u(x)$ of (5) members of which are made up of two exponential parts (where u is negative-valued) and two trigonometric parts (where u is positive-valued), and that these are the only two classes of oscillating solutions containing two exponential parts.*

By the same methods the following two theorems are obtained:

THEOREM VI. *There exist four classes of oscillating solutions $u(x)$ of (5) members of which are made up of three trigonometric parts and one exponential part.*

THEOREM VII. *There exists one class of oscillating solutions $u(x)$ of (5) members of which are made up wholly of trigonometric parts.*

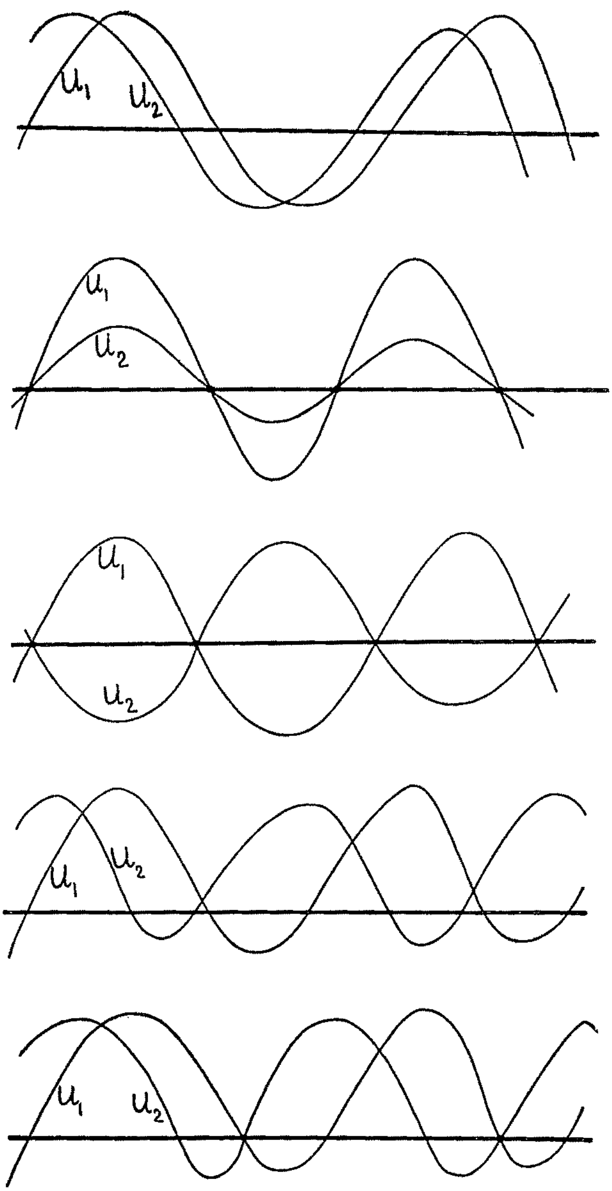
We list the possibilities and conditions for the sake of reference:

FOR OSCILLATING SOLUTIONS OF (5).

Type	Conditions	Number of Classes of Solutions
I. Four trigonometric parts.	$\Delta_1 < 0,\ \Delta_2 < 0,\ \Delta_3 < 0,\ \Delta_4 < 0.$	1.
II. Three trigonometric parts, One exponential part.	$\Delta_1 < 0,\ \Delta_2 < 0,\ \Delta_3 < 0,\ m_{4\nu}(+);$ $\Delta_1 < 0,\ \Delta_2 < 0,\ m_{3\nu}(-), \Delta_4 < 0;$ $\Delta_1 < 0,\ m_{2\nu}(+), \Delta_3 < 0,\ \Delta_4 < 0;$ $m_{1\nu}(-), \Delta_2 < 0,\ \Delta_3 < 0,\ \Delta_4 < 0.$	4.
III. Two trigonometric parts, Two exponential parts.	$\Delta_1 < 0,\ \Delta_2 < 0,\ m_{3\nu}(-), m_{4\nu}(+);$ $m_{1\nu}(-), m_{2\nu}(+), \Delta_3 < 0,\ \Delta_4 < 0.$	2.

Now in general the distance between consecutive zeros between which a solution is positive-valued will not be the same as the distance between consecutive zeros between which the solution is negative-valued. Whence it

follows that the zeros of two oscillating solutions u_1 and u_2 of (5) may [depending upon the coefficients a_1, a_2, b_1, b_1, c and upon the initial conditions $u_1(\gamma_1)=\delta_1$, $u_1'(\gamma_1)=\delta_1'$; $u_2(\gamma_2)=\delta_2$, $u_2'(\gamma_2)=\delta_2'$]: (1st), separate singly or coincide as in the theory of linear equations; (2nd), separate by pairs; and (3rd), as an intermediate case, alternately separate and coincide. (The figure explains the terminology employed).



Separation.

Coincidence.
(Linearly dependent solutions)

Coincidence.
(Linearly independent solutions)

Separation by pairs.

Alternate separation and coincidence.

4. *The Case of Variable Coefficients.* We begin the discussion by establishing the following

THEOREM VIII. *Consider the equation*

$$(1') \qquad u'' + p_1 u' + p_2 u + q_1 \mid u' \mid + q_2 \mid u \mid = 0$$

where p_1, p_2, q_1, q_2 *are real, single-valued and continuous in an interval* (a, b). *If neither* $(p_2 + q_2)$ *nor* $(p_2 - q_2)$ *is zero in* (a, b), *then between consecutive zeros of* u' *there is one* * *of* u.

Assume the contrary and consider first the case where u is positive and $(p_2 + q_2) > 0$. Now at points x_1, x_2 where $u' = 0$, our equation becomes $u'' = -(p_2 + q_2)u$ and u'' is negative; the curve is therefore concave downward. There must exist, however, a point $\bar{x}$ between x_1, x_2 where u' is zero because of the fact that u' is continuous and goes from minus to plus. But the differential equation is not satisfied by such a point for we have $(p_2 + q_2) > 0$, $u'(\bar{x}) = 0$, $u(\bar{x}) > 0$, and hence $u''(\bar{x})$ is positive. Similar arguments for the cases where u is positive and $(p_2 + q_2) < 0$, and where u is negative, $(p_2 - q_2) \gtrless 0$ will each lead to a contradiction from which the theorem follows. The corresponding theorem in the linear theory † appears as a special case when $q_1 \equiv q_2 \equiv 0$. Also it is of interest to note that if we remove the non-vanishing condition imposed on the quantities $(p_2 + q_2)$, $(p_2 - q_2)$, say from the latter, then the theorem will hold for u positive but not for u negative.

We shall now consider equation $(1')$ written in the form

$$(1'') \qquad d(Ku')/dx - L \mid u' \mid - N \mid u \mid - Gu = 0.$$

We shall have need to consider also the associated equations

$$(2'') \quad \begin{array}{ll} (2_1'') & d(Ku')/dx - Lu' - Nu - Gu = 0 \\ (2_2'') & d(Ku')/dx + Lu' - Nu - Gu = 0 \\ (2_3'') & d(Ku')/dx + Lu' + Nu - Gu = 0 \\ (2_4'') & d(Ku')/dx - Lu' + Nu - Gu = 0. \end{array}$$

Let u_1 be a solution of an equation of the form $(1'')$ with coefficients K_1, L_1, N_1, G_1; and similarly let u_2 be a solution of an equation of the form $(1'')$ with coefficients K_2, L_2, N_2, G_2; from which it follows that

* Rolle's Theorem would prohibit there being more than one.

† The author obtained this theorem from the lectures of R. G. D. Richardson, who has had it in his notes on "Linear Differential Equations of the Second Order" for years. The author is not aware of its presence elsewhere in the literature.

$$(11)\qquad d(K_1u_1')/dx - L_1\,|\,u_1'\,| - N_1\,|\,u_1\,| - G_1u_1 = 0$$
$$(12)\qquad d(K_2u_2')/dx - L_2\,|\,u_2'\,| - N_2\,|\,u_2\,| - G_2u_2 = 0.$$

Multiplying (11) by u_2 and (12) by $-u_1$ and adding we get, after simplifying by means of the equation themselves,

$$(13)\quad d(K_1u_1'u_2 - K_2u_2'u_1)/dx = (G_1 - G_2)u_1u_2 + (K_1 - K_2)u_1'u_2' + L_1\,|\,u_1'\,|\,u_2 - L_2\,|\,u_2'\,|\,u_1 + N_1\,|\,u_1\,|\,u_2 - N_2\,|\,u_2\,|\,u_1.$$

From this the following identity is readily obtained:

$$(14)\quad d[(u_1/u_2)(K_1u_1'u_2 - K_2u_2'u_1)]/dx = (G_1 - G_2)u_1^2 + (K_1 - K_2)u_1'^2 + K_2[(u_1'u_2 - u_2'u_1)/u_2]^2 + u_1^2[L_1(|\,u_1'\,|/u_1) - L_2(|\,u_2'\,|/u_2)] + u_1^2[N_1(|\,u_1\,|/u_1) - N_2(|\,u_2\,|/u_2)].$$

This is a generalization of the Picone identity * in the linear theory and reduces to that in the event $L \equiv N \equiv 0$.

Of two functions $y_1(x)$, $y_2(x)$, continuous in an interval (a, b), $y_2(x)$ will be said to oscillate more rapidly in (a, b) if it has more zeros than $y_1(x)$ in that interval. Let us now consider the question of the rapidity of the oscillation of the zeros of u_1 and u_2 (solutions of the two equations (11) and (12) respectively) under the following conditions on the coefficients: L_1, L_2, N_1, N_2 are non-negative; $K_1 \geq K_2 > 0$, $G_1 \geq G_2$; not all of the quantities L_1, L_2, N_1, N_2, G_1, G_2, are identically zero simultaneously. Let x_1, x_2 be consecutive zeros of u_1 which is taken to be positive between x_1, x_2. We can now state the following

THEOREM IX. *Any solution u_2 of* (12) *which is negative at some point γ: $x_1 \leq \gamma \leq x_2$, vanishes at least once in the open interval (x_1, x_2).*

Proof. Suppose, first, that $u_2(x_1) \neq 0$, $u_2(x_2) \neq 0$. Now $u_2(\gamma) < 0$ and $u_2(\xi)$, $x_1 \leq \xi \leq x_2$ will remain negative unless it vanishes in (x_1, x_2). Assuming that u_2 does not vanish, we integrate (14) between the limits x_1, x_2. The left-hand side yields zero. On the right-hand side we have a positive quantity in view of the conditions on the coefficients. The contradiction establishes the existence of at least one zero of u_2 in (x_1, x_2). Suppose, next, that u_2 vanishes at one or both of the end points of the interval considered. Even in this case it is true that u_2 has at least one zero in the open interval (x_1, x_2); for then at x_1 and at x_2 the determinate quantity u_1'/u_2' would

* See, for instance, Ince, "Ordinary Differential Equations," p. 226.

replace the indeterminate one u_1/u_2 in the left-hand member of (14) and an integration between the limits x_1, x_2 would give the desired contradiction. Thus, in (x_1, x_2), u_2 oscillates more rapidly than u_1.

If u_1 and u_2 are solutions of the same equation (1″), then equation (14) becomes, as a special case of considerable importance:

$$(15)\qquad d[(u_1/u_2)K(u_1'u_2 - u_2'u_1)]/dx = K[(u_1'u_2 - u_2'u_1)/u_2]^2 \\ + u_1^2L(|\,u_1'\,|/u_1 - |\,u_2'\,|/u_2) + u_1^2N(|\,u_1\,|/u_1 - |\,u_2\,|/u_2).$$

THEOREM X. *Given equation* (1″) *with the following conditions on the coefficients:* L, N *non-negative;* $K > 0$. *We consider a solution* u_1 *supposed positive between its consecutive zeros* x_1, x_2. *Then, a necessary and sufficient condition for another solution* u_2 *to vanish twice and only twice in the closed interval* (x_1, x_2) *is that the wronskian* $W(u_1, u_2)$ *be zero at some point* γ: $x_1 \leq \gamma \leq x_2$.

Proof of necessity. We first note that if u_2 is positive in (x_1, x_2) and vanishes at x_1 and x_2, then $W \equiv 0$ and the solutions are linearly dependent, and second that u_2 can not otherwise vanish twice in (x_1, x_2) and be positive between its zeros; for then the wronskian of two linearly independent solutions of one and the same associated linear equation [either $(2_1'')$ or $(2_2'')$] would have to be zero at some point in (x_1, x_2) and that is impossible. Hence if u_2 is to vanish twice in the interior of (x_1, x_2), it must be negative between its consecutive zeros. We have, then, u_2 vanishing twice in (x_1, x_2) and negative in between its two zeros, say γ_1, γ_2. If $\gamma_1 = x_1$, then $W[u_1(x_1), u_2(x_1)] = 0$. Similarly for $\gamma_2 = x_2$. And if $\gamma_1 = x_1$, $\gamma_2 = x_2$, x_1 and x_2 are simultaneous consecutive zeros of solutions u_1, u_2 which are in general linearly independent. This possibility we noted in the case of constant coefficients. Finally we take $\gamma_1 \neq x_1$, and $\gamma_2 \neq x_2$ with $u_2(\xi) < 0$, $\gamma_1 < \xi < \gamma_2$, and examine the ratios u_1'/u_1, u_2'/u_2. The former passes continuously through all real values from plus infinity to minus infinity between x_1, x_2 while the latter passes continuously through all real values from plus infinity to minus infinity in the sub-interval (γ_1, γ_2). These ratios are, consequently, equal at some point in (x_1, x_2) which implies that the wronskian vanishes. This completes the first part of the proof.

Proof of sufficiency. Let the wronskian be zero at some point γ. We have already seen that if u_2 is positive, then $W \equiv 0$ and the solutions are linearly dependent. Discarding this case, we see that u_2 is negative at the point γ and will remain negative throughout (x_1, x_2) unless it vanishes in that interval. We suppose that u_2 is negative throughout (x_1, x_2) and fix at-

tention on the identiy (15). Integrating (15) between the limits x_1 and γ, we have zero on the left-hand side and, because of the hypotheses on the coefficients, plus on the right-hand side which contradiction insures a zero of u_2 between x_1 and γ, say at γ_1. Similarly integrating between γ and x_2, we obtain another zero γ_2 of u_2. Finally, if γ_1 coincides with x_1 or if γ_2 coincides with x_2, an integration between x_1 and x_2 as limits will bring about the desired contradiction. That u_2 can have no zeros other than γ_1 and γ_2 we have already noted. The proof is therefore complete.

As immediate consequences of Theorem X, we set off the following corollaries:

COROLLARY I. *No solution u_2 can vanish twice in the open interval (x_1, x_2) and be positive between x_1, x_2.*

COROLLARY II. *If $W[u_1(\gamma), u_2(\gamma)] = 0$, $\gamma \neq x_1$, $\gamma \neq x_2$ and $W \not\equiv 0$, then $u_2(\gamma) < 0$ necessarily.*

COROLLARY III. *There exists a single infinitude of solutions u_2 each of which, in the interval of consecutive zeros x_1, x_2 of a given solution u_1, vanish twice and only twice.*

COROLLARY IV. *No solution u_2 can vanish more than twice in the closed interval (x_1, x_2); that is, separation is at most by pairs.*

The information gained from and the arguments used in the previous theorem are sufficient to establish the following

THEOREM XI. *Under the hypotheses of Theorem X, a necessary and sufficient condition for a solution u_2 to vanish once and only once in the open interval (x_1, x_2) is that the wronskian $W(u_1, u_2)$ be zero at no point in (x_1, x_2).**

Similar theorems could be developed for u_1 negative and u_2 positive by changing the conditions on the coefficients. The non-symmetric character of identities (14) and (15), however, suggests that it would be impossible to interchange the role of u_1 and u_2 and at the same time keep one and the same set of hypotheses on the coefficients. And it should be pointed out explicitly that, in the event separation is by pairs, say, as is the case in Theorem X, this separation by pairs does not necessarily persist either to the right or to the left of the interval (x_1, x_2). The same could be said about single separation. As an illustrative example, let us suppose that, in the interval (x_1, x_2), separation is by pairs u_2 vanishing at γ_1 and at γ_2 and fix our attention upon

* This becomes the separation theorem in the linear theory if $L \equiv N \equiv 0$.

a small interval Δx beginning at γ_2 and extending to the right. It is clear that the zeros of u_1 and u_2 will separate singly and mutually beyond Δx, if, in that interval, L, N pass to zero and remain identically zero to the right. As a matter of fact this is a more stringent condition than is necessary on the coefficients. Single separation would result to the right of Δx if the coefficients were changed continuously to constant values of such sort that there would result an equality of distance between consecutive zeros. And there would be other conditions on the coefficients which would necessarily bring about alternate separation of the zeros of u_1 and u_2.

Or it would be possible to pass from partial separation (i. e., alternate coincidence and separation) to single separation by the same type of changes on the coefficients. As a matter of fact, it is possible, in going from left to right, to pass from any one type of separation to any other with the exception, of course, of the case of coinciding zeros and linearly dependent solutions.

Frank Morley

The Problem of Lagrange in the Calculus of Variations.

By Gilbert Ames Bliss.

TABLE OF CONTENTS.

INTRODUCTION.

The problem of the calculus of variations principally considered in this paper is that of finding in a class of arcs

$$(1) \qquad y_i = y_i(x) \qquad (x_1 \leqq x \leqq x_2;\ i = 1, \cdots, n)$$

satisfying a set of differential equations

$$(2) \qquad \phi_\alpha(x, y_1, \cdots, y_n, y_1', \cdots, y_n') = 0 \qquad (\alpha = 1, \cdots, m < n)$$

and joining two fixed points in the space of points $(x, y_1, \cdots, y_n)$, one which minimizes an integral of the form

$$(3) \qquad I = \int_{x_1}^{x_2} f(x, y_1, \cdots, y_n, y_1', \cdots, y_n')\, dx.$$

A number of paragraphs are also devoted to the similar problem for which the end-points are variable.

The problem seems to have been first formulated by Lagrange for the general case here studied, though somewhat less precisely than in the statement above. He also gave the multiplier rule described in Section 5 below which had been previously deduced by Euler and himself for a number of more special cases. Important additions to the theory have been made by Clebsch, A. Mayer, Kneser, Hilbert, von Escherich, Hahn, Bolza, and many others. Comprehensive treatments of the problem have been given by Bolza [3] * and Hadamard [4], that of Bolza being the more complete. In Chapter V below a brief sketch of the history of the problem is given with a bibliography of the more important papers on which the text of this paper is based.

Since the literature of the problem is extensive and widely scattered, and since recent developments make possible important simplifications, even as compared with the excellent treatments of Bolza and Hadamard, it seemed justifiable to the author of this paper to attempt anew the presentation of those parts of the theory leading to the necessary conditions for a minimum, and to those sufficient to insure a minimum. The paper is a record of lectures which the author has given at intervals for some years past at the University of Chicago.

Some special features of the methods used may perhaps be mentioned. The deduction of the Euler-Lagrange multiplier rule in Sections 3-5 is based upon suggestions in papers by Hahn [13, p. 271] and the author [16, pp. 307, 312], but is different from the proofs hitherto given. The definition of

* The figures in the square brackets refer to the bibliographical list at the end of the paper.

normal arcs in Sections 7 and 8 is that of Bolza [19, p. 440]. A new application of the definition, in Section 15, makes it possible to deduce without the use of special methods the multiplier rule for the case when the functions ϕ_α contain none of the derivatives y_i', as a corollary to the rule deduced in Section 5. The discussions of the necessary conditions of Weierstrass and Clebsch, and of the envelope theorem with the associated deduction of the necessary condition of Mayer, are essentially those of Hahn [21] and Bolza [3, pp. 603-10], but are greatly simplified by the use of the auxiliary formulas of Section 21. The analytic proof of the necessary condition of Mayer in Section 26, by means of the minimum problem associated with the second variation, was suggested by the author for simpler cases [27] and applied to the problem of Lagrange by D. M. Smith [28]. By means of the theory of the minimum problem of the second variation the very elaborate theories of that variation due to Clebsch [29], von Escherich [31], Hahn [33], and others, can be much simplified, as the author has shown [35]. The applications important for this paper are in Sections 26 and 32. The theory of Mayer fields in Sections 28 and 29, and the proofs of the sufficiency theorems in Sections 30 and 31, have been simplified as far as seemed possible.

An effort has been made in each theorem to state clearly the underlying hypotheses. The proof of the multiplier rule in Section 5, for example, is independent of the assumption that the determinant R of page 11 is different from zero. In many of the succeeding theorems, however, this assumption is either made explicitly or else is a consequence of the property III′ which appears frequently.

CHAPTER I.

THE EULER-LAGRANGE MULTIPLIER RULE.

1. *Hypotheses.* In this first chapter the famous multiplier rule of Euler and Lagrange, describing the differential equations satisfied by a minimizing arc for the problem of Lagrange stated in the introduction, is to be deduced. For convenience in the following pages the set $(x, y_1, \cdots, y_n, y_1, \cdots, y_n')$ will be represented by (x, y, y').

As usual we concentrate attention on a particular arc E_{12} with the equations (1) and inquire what properties it must have if it is to be a minimizing arc. The analysis is based upon the following hypotheses:

(a) the functions $y_i(x)$ defining E_{12} are continuous on the interval x_1x_2 and this interval can be subdivided into a finite number of parts on each of which the functions have continuous derivatives;

(b) in a neighborhood $\mathfrak{R}$ of the values (x, y, y') on the arc E_{12} the functions f, ϕ_α have continuous derivatives up to and including those of the **fourth order;**

(c) at every element (x, y, y') on E_{12} the $m \times n$-dimensional matrix $\| \phi_{\alpha y_i'} \|$ has rank m.

The subscript y_i' here indicates the partial derivative of ϕ_α with respect to y_i'. In the following pages literal subscripts, following the indices of functions and elsewhere, will be frequently used to indicate partial derivatives. The hypothesis (c) implies that the equations $\phi_\alpha = 0$ are all independent near E_{12} when regarded as functions of the variables y_i'.

2. *Examples.* A common example of a Lagrange problem is that of the brachistochrone in a resisting medium [3, p. 5]. The differential equation of the motion [5, p. 44] becomes for this case

$$dv/dt = d^2s/dt^2 = g(dy/ds) - R(v),$$

where $R(v)$ is the retardation on the particle per unit mass due to the resistance of the medium. Multiplying by $ds/dx = (ds/dt)(dt/dx) = v dt/dx$ we find the equation

$$vv' = gy' - R(v)s' = gy' - R(v)(1 + y'^2)^{1/2} \tag{4}$$

where the primes denote derivatives with respect to x. The problem is then to find among the pairs of functions $y(x)$, $v(x)$ which have the end-values

$$y(x_1) = y_1, \quad v(x_1) = v_1, \quad y(x_2) = y_2$$

and satisfy equation (4) one which minimizes the time integral

$$I = \int_{s_1}^{s_2} (ds/v) = \int_{x_1}^{x_2} (1/v)(1 + y'^2)^{1/2} dx.$$

It should be noted that this problem is not precisely like that stated in section 1 since the value of v is not prescribed at $x = x_2$. It is in fact a problem of Lagrange with second end-point variable.

The so-called isoperimetric problems form a very large class, and all of them may be stated as Lagrange problems. For example we may seek to find among the arcs $y = y(x)$ $(x_1 \leqq x \leqq x_2)$, joining two given points and having a given length, one which has its center of gravity the lowest. This is the problem of determining the form of a hanging chain suspended between two pegs at its ends. Analytically the problem is to find among the functions $y(x)$ $(x_1 \leqq x \leqq x_2)$ satisfying the conditions

$$y(x_1)=y_1, \quad y(x_2)=y_2, \qquad \int_{x_1}^{x_2}(1+y'^2)^{1/2}dx=l$$

one which minimizes the integral

$$I=\int_{x_1}^{x_2} y(1+y'^2)^{1/2}dx. \tag{5}$$

This problem may be made over into one of the Lagrange type by introducing the new variable

$$z(x)=\int_{x_1}^{x}(1+y'^2)^{1/2}dx$$

satisfying the differential equation $z'=(1+y'^2)^{1/2}$. The problem is then to find among the pairs $y(x)$, $z(x)$ satisfying $y(x_1)=y_1$, $z(x_1)=0$, $y(x_2)=y_2$, $z(x_2)=l$, $z'=(1+y'^2)^{1/2}$ one which minimizes the integral (5).

More generally suppose we wish to find among the functions $y(x)$ satisfying

$$y(x_1)=y_1, \quad y(x_2)=y_2$$

$$\int_{x_1}^{x_2} g(x,y,y')\,dx=k, \qquad \int_{x_1}^{x_2} h(x,y,y')\,dx=l$$

one which minimizes

$$I=\int_{x_1}^{x_2} f(x,y,y')\,dx. \tag{6}$$

The problem is equivalent to that of finding among the sets of functions $y(x)$, $u(x)$, $v(x)$ satisfying

$$\begin{aligned} &y(x_1)=y_1, \quad u(x_1)=0, \quad v(x_1)=0,\\ &y(x_2)=y_2, \quad u(x_2)=k, \quad v(x_2)=l,\\ &u'=g(x,y,y'), \qquad v'=h(x,y,y'), \end{aligned}$$

one which minimizes the integral (6). Evidently a similar transformation of the problem could be made no matter how many isoperimetric integrals were to have prescribed constant values.

These illustrations suffice to show the wide applicability of the Lagrange problem.

3. *Admissible arcs and variations.* An *admissible arc*

$$y_i=y_i(x) \qquad (i=1,\cdots,n;\ x_1\leqq x\leqq x_2) \tag{7}$$

is one with the continuity properties (a) of Section 1, whose elements (x,y,y') all lie in the region $\mathfrak{R}$, and which satisfies the equations $\phi_\alpha=0$. If a one-parameter family of admissible arcs

$$(8) \qquad y_i = y_i(x, b) \qquad (i = 1, \cdots, n)$$

containing a particular admissible arc E_{12} for the parameter value $b = b_0$ is given, the functions

$$\eta_i(x) = y_{ib}(x, b_0) \qquad (i = 1, \cdots, n),$$

where the subscript b indicates as usual a partial derivative of $y_i(x, b)$, are called variations of the family along E_{12}.

In the tensor analysis it is agreed that a product $G_{ik}H_k$ shall stand for the sum $\Sigma_k G_{ik}H_k$. In other words, when an index k occurs twice in the same term it is understood that the term really represents the sum of n terms of the same type. The index with respect to which the sum is taken is called an umbral index.

With this convention in mind we may define for the arc E_{12} mentioned above the so-called *equations of variation* by the formula

$$(9) \qquad \Phi_\alpha(x, \eta, \eta') = \phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i' = 0 \qquad (\alpha = 1, \cdots, m)$$

in which i is an umbral index with the range $1, \cdots, n$, and the coefficients $\phi_{\alpha y_i}$, $\phi_{\alpha y_i'}$ are supposed to have as arguments the functions $y_i(x)$ belonging to E_{12}. These equations are satisfied by the variations $\eta_i(x)$ along E_{12} as we may readily see by substituting the functions (8) in the equations $\phi_\alpha = 0$, differentiating for b, and setting $b = b_0$. A set of functions $\eta_i(x)$ with the continuity properties described in (a) of Section 1 and satisfying the equations of variation (9) is called a *set of admissible variations,* a nomenclature which is justified by the following very important theorem:

For every set of admissible variations $\eta_i(x)$ along the admissible arc E_{12} there exists a one parameter family (8) of admissible arcs containing E_{12} for the value $b = 0$ and having the functions $\eta_i(x)$ as its variations along E_{12}. For this family the functions $y_i(x, b)$ are continuous and have continuous derivatives with respect to b for all values (x, b) near those defining E_{12}, and the derivatives $y_{ix}(x, b)$ have the same property except possibly at the values of x defining corners of E_{12}.

To prove this theorem we enlarge the system $\phi_\alpha = 0$ to have the form

$$(10) \qquad \phi_1 = 0, \cdots, \phi_m = 0, \quad \phi_{m+1} = z_{m+1}, \cdots, \phi_n = z_n$$

where $z_{m+1}, \cdots, z_n$ are new variables and $\phi_{m+1}, \cdots, \phi_n$ are new functions of x, y, y' such that the functional determinant $|\partial\phi_i/\partial y_k'|$ is different from zero along E_{12}.* By means of the last $n - m$ of these equations the functions

* For a proof of the possibility of this adjunction see Bliss [16, pp. 307, 312].

$y_i(x)$ belonging to E_{12} define a set of functions $z_r(x)$ $(r = m+1, \cdots, n)$. We have a corresponding system of equations of variation

$$(11) \qquad \Phi_1 = 0, \cdots, \Phi_m = 0, \quad \Phi_{m+1} = \zeta_{m+1}, \cdots, \Phi_n = \zeta_n$$

along E_{12}, the last $n - m$ of which define a set $\zeta_r(x)$ $(r = m+1, \cdots, n)$ corresponding to every set of admissible variations $\eta_i(x)$.

Suppose now that the set $\eta_i(x)$ is an admissible set of variations for E_{12} defining a set $\zeta_r(x)$ by means of equations (11). Since the functional determinant $|\partial\phi_i/\partial y_k'|$ is different from zero along E_{12} the existence theorems for differential equations * tell us that the system

$$(12) \quad \phi_\alpha = 0, \quad \phi_r = z_r(x) + b\zeta_r(x) \quad (\alpha = 1, \cdots, m;\ r = m+1, \cdots, n)$$

determines uniquely a one-parameter family of solutions $y_i = y_i(x, b)$ with the initial values $y_i(x_1) + b\eta_i(x_1)$ at $x = x_1$. This family contains E_{12} for $b = 0$ and has variations which have the initial values $\eta_i(x_1)$ at $x = x_1$ and which satisfy the equations (11) with the functions $\zeta_r(x)$. The variations of the family are therefore identical with the functions $\eta_i(x)$ originally prescribed, since when the $\zeta_r(x)$ are given there is only one set of solutions of equations (11) with given initial values $\eta_i(x_1)$ at $x = x_1$.

Some slight modifications in the existence theorems referred to are required in order to prove the continuity properties of the family $y_i = y_i(x, b)$ described in the theorem. These are due to the fact that the functions $z_i(x)$ defined by the arc E_{12} are continuous but not necessarily differentiable. The results described can be derived without difficulty, however, when the arc E_{12} has no corners. If the arc E_{12} has corners the existence theorems must be applied successively to the x-intervals between the corner-values of x with initial conditions at the beginning of each interval so chosen that the functions $y_i(x, b)$ are continuous.

COROLLARY. *If a matrix*

$$\begin{Vmatrix} \eta_{11} & \cdot & \cdot & \cdot & \eta_{1\mu} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1} & \cdot & \cdot & \cdot & \eta_{n\mu} \end{Vmatrix}$$

whose columns are μ sets of admissible variations along an admissible arc E_{12}, is given, then there exists a μ-parameter family of admissible arcs $y_i = y_i(x, b_1, \cdots, b_\mu)$ containing E_{12} for the values $b_1 = \cdots = b_\mu = 0$ and having the functions η_{is} $(i = 1, \cdots, n)$ as its variations with respect to b_s along E_{12}. The continuity properties of the family are similar to those described in the preceding theorem.

---

* Bolza [3, pp. 168 ff.]; Bliss [14, 15].

This is proved as above with the equations

$$\phi_\alpha = 0, \quad \phi_r = z_r(x) + b_1\zeta_{r1} + \cdots + b_\mu\zeta_{r\mu}$$
$$(\alpha = 1, \cdots, m;\ r = m+1, \cdots, n)$$

replacing equations (12).

4. *The first variation of* I. If the functions $y_i(x, b)$ defining a one-parameter family of admissible arcs containing E_{12} for $b = 0$ are substituted in I then I becomes the function of b' defined by the formula

$$I(b) = \int_{x_1}^{x_2} f[x, y(x, b), y'(x, b)]\,dx.$$

The derivative of this function with respect to b at the value $b = 0$ is the expression

$$(13) \qquad I_1(\eta) = \int_{x_1}^{x_2} (f_{y_i}\eta_i + f_{y_i'}\eta_i')\,dx$$

where i is as agreed an umbral symbol and the arguments of the derivatives of f are the functions $y_i(x)$ defining E_{12}.

The expression $I_1(\eta)$ is called the *first variation* of I along the arc E_{12}. For the proofs of the succeeding sections it is desirable to have another form of it. Let λ_0 be a constant and $\lambda_i(x)$ $(i = 1, \cdots, n)$ functions of x on the interval x_1x_2, and let F be defined by the equation

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1\phi_1 + \cdots + \lambda_n\phi_n.$$

Since the variations η, ζ satisfy the equations (11) the value of $\lambda_0 I_1(\eta)$ is not altered if we add the sum $\lambda_\alpha\Phi_\alpha + \lambda_r(\Phi_r - \zeta_r)$ to its integrand. Then we have

$$(14) \qquad \lambda_0 I_1(\eta) = \int_{x_1}^{x_2} (F_{y_i}\eta_i + F_{y_i'}\eta_i' - \lambda_r\zeta_r)\,dx.$$

So far the functions $\lambda_i(x)$ have been entirely arbitrary. We now determine them so that the equations

$$(15) \qquad F_{y_i'} = \int_{x_1}^{x} F_{y_i}\,dx + c_i \qquad (i = 1, \cdots, n)$$

are satisfied for an arbitrarily selected set of constants λ_0, c_i. This is possible since if we introduce the new variables

$$(16) \qquad v_i = F_{y_i'} = \lambda_0 f_{y_i'} + \lambda_1\phi_{1y_i'} + \cdots + \lambda_n\phi_{ny_i'} \qquad (i = 1, \cdots, n)$$

the equations (15) are equivalent to the equations and initial conditions

$$(17) \qquad dv_i/dx = F_{y_i} = A_{i1}v_1 + \cdots + A_{in}v_n + B_i, \quad v_i(x_1) = c_i \qquad (i = 1, \cdots, n)$$

the coefficients A, B being found by solving the equations (16) for $\lambda_1, \cdots, \lambda_n$ and substituting in F_{y_i}. The equations (17) have unique solutions $v_i(x)$ which are continuous on the interval x_1x_2 and which have continuous derivatives except possibly at the values of x defining the corners of E_{12} where the coefficients A, B may be discontinuous. Equations (16) then determine uniquely the functions $\lambda_i(x)$ continuous except possibly at the corner values of x.

With the help of equations (15) the expression (14) for $\lambda_0 I_1(\eta)$ now takes the form

$$\lambda_0 I_1(\eta) = -\int_{x_1}^{x_2} \lambda_r \zeta_r dx - c_i \eta_i(x_1) + \eta_i(x_2) F_{y_i'}(x_2) \tag{18}$$

where $F_{y_i'}(x_2)$ represents the value of $F_{y_i'}$ at $x = x_2$. This auxiliary formula will be useful in the next section.

5. *The Euler-Lagrange multiplier rule.* We are now in a position to deduce the famous multiplier rule giving the differential equations which must be satisfied by a minimizing arc E_{12} for the Lagrange problem. The rule was discussed for a special case by Euler in 1744, and generalized by Lagrange whose proof was exceedingly faulty. One difficulty with Lagrange's proof was overcome by Mayer in 1886, and the proof was finally completed when Kneser in 1900 and Hilbert in 1905 removed the last serious defects.* The proof given here is quite different in some respects from those in the literature and is an extension of them.

Suppose that a matrix whose columns are $2n + 1$ sets of admissible variations

$$\left\| \begin{matrix} \eta_{11} & \cdot & \cdot & \cdot & \eta_{1,2n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1} & \cdot & \cdot & \cdot & \eta_{n,2n+1} \end{matrix} \right\| \tag{19}$$

is given. We have seen above that there is a $(2n + 1)$-parameter family $y_i(x, b_1, \cdots, b_{2n+1})$ of admissible arcs containing E_{12} for $b_1 = \cdots = b_{2n+1} = 0$ and having the columns of the matrix above as its variations. When the functions defining this family are inserted in the integral I that integral becomes a function $I(b_1, \cdots, b_{2n+1})$ which for $b_1 = \cdots = b_{2n+1} = 0$ takes the value I_0 of the integral along the arc E_{12}. If we let $(x_1, y_{11}, \cdots, y_{n1})$ and $(x_2, y_{12}, \cdots, y_{n2})$ represent the two end-points on the arc E_{12} then the equations

* For the details of the objections to Lagrange's proof and an excellent historical sketch see Bolza [3, p. 566].

$$
(20) \qquad \begin{gathered} I(b_1, \cdots, b_{2n+1}) = I_0 + u, \\ y_i(x_1, b_1, \cdots, b_{2n+1}) = y_{i1}, \\ y_i(x_2, b_1, \cdots, b_{2n+1}) = y_{i2}, \\ (i = 1, \cdots, n) \end{gathered}
$$

in the variables $u, b_1, \cdots, b_{2n+1}$ have the initial solution $(u, b_1, \cdots, b_{2n+1}) = (0, 0, \cdots, 0)$. If the functional determinant of the first members of these equations with respect to $b_1, \cdots, b_{2n+1}$ is different from zero at this solution, then well-known implicit function theorems tell us that the equations (20) have solutions not only for $u = 0$ but also for every value of u near $u = 0$. There are therefore arcs in the family $y_i(x, b_1, \cdots, b_{2n+1})$ joining the end-points 1 and 2 of E_{12} and giving I values $I_0 + u$ greater than I_0 when u is positive, and similar arcs giving it values less than I_0 when u is negative, which is impossible if E_{12} is a minimizing arc. Hence the functional determinant of the equations (20) must be zero at $(u, b_1, \cdots, b_{2n+1}) = (0, 0, \cdots, 0)$.

The value of this functional determinant is

$$
(21) \qquad \begin{Vmatrix} I_1(\eta_1) & \cdots & I_1(\eta_{2n+1}) \\ \eta_{11}(x_1) & \cdots & \eta_{1,2n+1}(x_1) \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \eta_{n1}(x_1) & \cdots & \eta_{n,2n+1}(x_1) \\ \eta_{11}(x_2) & \cdots & \eta_{1,2n+1}(x_2) \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \eta_{n1}(x_2) & \cdots & \eta_{n,2n+1}(x_2) \end{Vmatrix}
$$

where in the first row only the second subscripts of the η's are indicated. It must vanish for every choice of the matrix (19) of admissible variations. Suppose $p < 2n + 1$ the highest rank attainable for (21) and suppose the matrix (19) chosen so that this rank is actually attained. Let λ_0, c_i, d_i $(i = 1, \cdots, n)$ be a set of constants not all zero satisfying the linear equations whose coefficients are the columns of the determinant (21). Normally the constant λ_0 will be different from zero, but in Section 7 the case $\lambda_0 = 0$ is discussed in more detail. In both cases the equation

$$
\lambda_0 I_1(\eta) + c_i\eta_i(x_1) + d_i\eta_i(x_2) = 0
$$

must be satisfied for every set of admissible variations $\eta_i(x)$ whatsoever, since otherwise by deleting a suitable one of the columns of the determinant (21) and replacing it by a set $I_1(\eta)$, $\eta_i(x_1)$, $\eta_i(x_2)$ which does not satisfy the last equation, the determinant could be made to have the rank $p + 1$. If the first term of the last equation is replaced by its value (18) the equation takes the form

$$-\int_{x_1}^{x_2} \lambda_r \zeta_r dx + \eta_i(x_2)[d_i + F_{y_i'}(x_2)] = 0$$

and it must be satisfied for every choice of the admissible variations $\eta_i(x)$, i. e. for every choice of the functions $\zeta_r(x)$ and the end values $\eta_i(x_2)$, since for every such choice there is a set of admissible variations defined by the equations (11). It follows readily that the conditions

$$\lambda_r(x) \equiv 0, \quad d_i = -F_{y_i'}(x_2) \qquad (r = m+1, \cdots, n;\ i = 1, \cdots, n) \tag{22}$$

must be satisfied. For the set of multipliers λ_0, $\lambda_i(x)$ $(i = 1, \cdots, n)$ for which the equations (15) are satisfied it is evident then that all are identically zero except the first $m+1$. The first $m+1$ of them are not all identically zero, however, since otherwise F would vanish identically and equations (15) and (22) would require the constants c_i, d_i all to be zero as well as λ_0, which we know not to be the case. Hence we have the following theorem:

For every minimizing arc E_{12} there exists a set of constants c_i $(i = 1, \cdots, n)$ and a function

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1(x)\phi_1 + \cdots + \lambda_m(x)\phi_m \tag{23}$$

such that the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i \tag{24}$$

are satisfied at every point of E_{12}. The constant λ_0 and the functions $\lambda_\alpha(x)$ $(\alpha = 1, \cdots, m)$ are not all identically zero on x_1x_2 and are continuous except possibly at values of x defining corners of E_{12}.

This is a modification of the Euler-Lagrange multiplier rule. We get the rule in its classical form by differentiating the equations (24). The two following corollaries are immediate:

COROLLARY I. THE EULER-LAGRANGE MULTIPLIER RULE. *On every sub-arc between corners of a minimizing arc E_{12} the differential equations*

$$\phi_\alpha(x, y, y') = 0, \quad (d/dx)F_{y_i'} = F_{y_i} \quad (\alpha = 1, \cdots, m;\ i = 1, \cdots, n) \tag{25}$$

must be satisfied, where F is the function (23).

COROLLARY II. THE CORNER CONDITION. *At every corner of a minimizing arc E_{12} the conditions*

$$F_{y_i'}[x, y, y'(x-0), \lambda(x-0)] = F_{y_i'}[x, y, y'(x+0), \lambda(x+0)] \qquad (i = 1, \cdots, n) \tag{26}$$

must be satisfied.

Condition (26) is a consequence of the fact that the second member of (24) is continuous at a corner as well as elsewhere.

There is a third consequence of the equations (24) which is also important. If the functions and multipliers belonging to E_{12} are $y_i(x)$, λ_0, $\lambda_a(x)$ then the $n+m$ equations

$$F_{y_i'}[x, y(x), z, \mu] = \int_{x_1}^{x} F_{y_i}[x, y(x), y'(x), \lambda(x)]dx + c_i,$$

$$\phi_a[x, y(x), z] = 0 \qquad (i = 1, \cdots, n;\ \alpha = 1, \cdots, m)$$

have as solutions the $n+m$ functions $z_i = y_i'(x)$, $\mu_a = \lambda_a(x)$. If the functional determinant

$$R = \begin{vmatrix} F_{y_i' y_k'} & \phi_{a y_i'} \\ \phi_{a y_k'} & 0 \end{vmatrix}$$

of the first members of these equations with respect to the variables z_i, μ_a is different from zero at a point of E_{12} then the existence theorems for implicit functions tell us that the solutions $z_i = y_i'(x)$, $\mu_a = \lambda_a(x)$ of the equations have continuous derivatives of as many orders as the equations themselves have continuous partial derivatives in the variables x, z_i, μ_a. Between corners this is at least one, and we have the following third corollary:

COROLLARY III. THE DIFFERENTIABILITY CONDITION. *Near a point of a minimizing arc E_{12} at which the determinant R is different from zero the functions $y_i(x)$ defining E_{12} have continuous second derivatives and the multipliers $\lambda_a(x)$ have continuous first derivatives.*

The proof given above for the Euler-Lagrange multiplier rule is an extension of the ones ordinarily given because the hypothesis (c) Section 1 is less restrictive than usual. The unsymmetrical assumption commonly made is that a particular one of the determinants of the matrix $\| \phi_{a y_i'} \|$ stays different from zero at every point of E_{12}. The enlargement of the system $\phi_a = 0$ to the system (10) is the device which permits the generalization here made. Equations (24) are recent developments which were unknown to Euler and Lagrange and which are not always deduced even in modern presentations of the subject. They justify the useful Corollaries II and III besides the multiplier rule.

6. *The extremals.* An admissible arc and set of multipliers

$$y_i = y_i(x), \quad \lambda_0, \quad \lambda_a = \lambda_a(x) \qquad (27)$$
$$(i = 1, \cdots, n;\ \alpha = 1, \cdots, m;\ x_1 \leqq x \leqq x_2)$$

is called an *extremal* if it has continuous derivatives $y_i'(x)$, $y_i''(x)$, $\lambda_a'(x)$

on the interval x_1x_2, and if furthermore it satisfies the Euler-Lagrange equations (25). The minimizing curves for applications of the theory of the calculus of variations are found among the extremals and it is highly desirable, therefore, that we should examine more thoroughly the differential equations defining these curves and determine how large a family the extremals really form. A minimizing curve must always be a solution of the equations (25), even if it has corners or is without the derivatives $y_i''(x)$, $\lambda_a'(x)$ mentioned above, but such minimizing curves are relatively rare.

The most direct way to characterize the family of extremals satisfying equations (25) is to replace these equations by the equivalent system

$$
(28) \qquad \begin{aligned}
(d/dx)F_{y_i'} - F_{y_i} = F_{y_i'x} + F_{y_i'y_k}y_k' + F_{y_i'y_k'}y_k'' + F_{y_i'\lambda_\beta}\lambda_\beta' - F_{y_i} = 0,\\
(d/dx)\phi_a = \phi_{ax} + \phi_{ay_k}y_k' + \phi_{ay_k'}y_k'' = 0,\\
\phi_a\,[x_1, y(x_1), y'(x_1)] = 0.
\end{aligned}
$$

The first two of these equations are linear in the variables y_k'', λ_β' and the determinant of the coefficients of these variables is precisely the determinant R of page 684. Near an extremal E_{12} on which R is different from zero these two equations can therefore be solved for y_k'', λ_β' and they are readily seen to be equivalent to a system

$$
(29) \qquad dy_k/dx = y_k', \quad dy_k'/dx = G_k(x, y, y', \lambda), \quad d\lambda_\beta/dx = H_\beta(x, y, y', \lambda)
$$

in the so-called normal form.* Known existence theorems for differential equations now tell us that an extremal E_{12} along which R is different from zero is a member of a family of solutions of equations (29) depending upon $2n + m$ arbitrary constants, since the number of dependent variables y_k, y_k', λ_β in these equations is $2n + m$. If we impose further the m relations in the third row of equations (28) then m of these constants will be determined as function of the $2n$ others, so that the final result is that an extremal along which R is different from zero is a member of a $2n$-parameter family of extremals satisfying equations (25).

For theoretical purposes the properties of the $2n$-parameter family of extremals may be determined most conveniently by a second method.† For the purpose of introducing n new variables v_i and eliminating the $n + m$ variables y_i', λ_a let us consider the system of $n + m$ equations

$$
(30) \qquad F_{y_i'}(x, y, y', \lambda) = v_i, \quad \phi_a(x, y, y') = 0.
$$

The functional determinant of the first members of these equations with respect

* Bolza [3, p. 589].

† Bolza [3, p. 590].

to the variables y_k', λ_β is again the determinant R of page 684. Known theorems on implicit functions tell us then that near an extremal E_{12} on which R is different from zero the equations (30) have solutions

$$(31) \qquad y_k' = \Psi_k(x, y, v), \quad \lambda_\beta = \Pi_\beta(x, y, v)$$

possessing continuous partial derivatives of the first three orders since the first members of equations (30) have such derivatives. The system of equations (25) is now equivalent to the system in normal form

$$(32) \quad dy_k/dx = \Psi_k(x, y, v), \quad dv_k/dx = F_{y_i}[x, y, \Psi(x, y, v), \quad \Pi(x, y, v)]$$

in the variables x, y_k, v_k. Evidently every solution $y_k(x)$, $\lambda_\beta(x)$ of equations (25) defines a set of functions $v_k(x)$ satisfying equations (30) and (31), and therefore also the system (32). Conversely every solution $y_k(x)$, $v_k(x)$ of equations (32) defines a set of functions $\lambda_\beta(x)$ by means of equations (31) with which it satisfies equations (30), and therefore also the original system (25).

Through every initial element

$$(x_0, y_0, v_0) = (x_0, y_{10}, \cdots, y_{n0}, v_{10}, \cdots, v_{n0})$$

in a neighborhood of the set of values (x, y, v) on the extremal E_{12} there passes a unique solution

$$(33) \qquad y_i = y_i(x, x_0, y_0, v_0), \quad v_i = v_i(x, x_0, y_0, v_0)$$

of the equations (32) for which the functions y_i, y_{ix}, v_i, v_{ix} have continuous partial derivatives of the first three orders since the second members of equations (32) have such derivatives. The equations expressing the fact that the solutions (33) passes through (x_0, y_0, v_0) are

$$y_{i0} = y_i(x_0, x_0, y_0, v_0), \quad v_{i0} = v_i(x_0, x_0, y_0, v_0),$$

and from them we find

$$(34) \qquad \begin{aligned} \delta_{ik} &= (\partial/\partial y_{k0})\, y_i(x_0, x_0, y_0, v_0), & 0 &= (\partial/\partial v_{k0})\, y_i(x_0, x_0, y_0, v_0), \\ 0 &= (\partial/\partial y_{k0})\, v_i(x_0, x_0, y_0, v_0), & \delta_{ik} &= (\partial/\partial v_{k0})\, v_i(x_0, x_0, y_0, v_0), \end{aligned}$$

where δ_{ik} is 1 or 0 when $k = i$ or $k \neq i$, respectively. Since every curve of this system (33) has on it an initial element for which $x = x_1$ we lose none of the curves if we replace x_0 by the fixed value x_1. Let us for convenience rename the constants y_{i0}, v_{i0} and call them a_i, b_i respectively. Then the family (33) takes the form

$$(35) \qquad y_i = y_i(x, a, b), \quad v_i = v_i(x, a, b)$$

and it follows readily from equations (34) that the determinant

$$(36)\qquad \begin{vmatrix} \dfrac{\partial y_i}{\partial a_k} & \dfrac{\partial y_i}{\partial b_k} \\ \dfrac{\partial v_i}{\partial a_k} & \dfrac{\partial v_i}{\partial b_k} \end{vmatrix}$$

has the value 1 at $x = x_1$. When we substitute the functions (35) in equations (31) a set of functions $\lambda_a(x, a, b)$ is determined, and we have the final result:

Every extremal E_{12} along which the determinant R is different from zero is a member of a $2n$-parameter family of extremals

$$(37)\qquad y_i = y_i(x, a, b), \quad \lambda_a = \lambda_a(x, a, b)$$

for special values a_0, b_0 of the parameters. The functions y_i, y_{ix}, v_i, v_{ix}, λ_β have continuous partial derivatives of the first three orders in a neighborhood of the values (x, a, b) defining E_{12}, and at the special values (x_1, a_0, b_0) the determinant (36) *is different from zero.*

Thus again we have established the existence of a family of extremals containing $2n$ arbitrary constants.

7. *Normal admissible arcs.* An admissible arc $y_i = y_i(x)$ $(x_1 \leqq x \leqq x_2)$ is said to be *normal* if there exist for it $2n$ sets of admissible variations for which the determinant

$$(38)\qquad \begin{vmatrix} \eta_{11}(x_1) & \cdots & \eta_{1,2n}(x_1) \\ \cdot & \cdots & \cdot \\ \eta_{n1}(x_1) & \cdots & \eta_{n,2n}(x_1) \\ \eta_{11}(x_2) & \cdots & \eta_{1,2n}(x_2) \\ \cdot & \cdots & \cdot \\ \eta_{n1}(x_2) & \cdots & \eta_{n,2n}(x_2) \end{vmatrix}$$

is different from zero. It is *normal on a sub-interval* $\xi_1\xi_2$ of x_1x_2 if there exist $2n$ sets of admissible variations for which the last determinant is different from zero when x_1 is replaced by ξ_1 and x_2 by ξ_2. In the sequel we shall frequently need to restrict our proofs to arcs which are *normal on every sub-interval of* x_1x_2.

These definitions doubtless seem at first sight somewhat artificial. If an admissible arc E_{12} is not normal, however, it is in general true that no other admissible arcs near it pass through the end points 1 and 2 of E_{12}, and hence that near E_{12} the class of arcs in which we seek to minimize the integral I has in it only E_{12} itself. The minimum problem in such a case would not be

of interest. We shall presently see that there are always an infinity of admissible arcs through the ends of E_{12} when E_{12} is normal.

A necessary and sufficient condition that an admissible arc be normal is that there exists for it no set of multipliers λ_0, $\lambda_a(x)$ *having* $\lambda_0 = 0$ *with which it satisfies the equations.*

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i.$$

For a normal extremal arc multipliers in the form $\lambda_0 = 1$, $\lambda_a(x)$ *always exist and in this form they are unique.*

The processes of Section 5 show that an admissible arc which is not normal has surely a set of multipliers with $\lambda_0 = 0$, since the linear equations whose coefficients are the columns of the determinant (21) have for such an arc a set of solutions λ_0, c_i, d_i with $\lambda_0 = 0$. The first sentence of the theorem will then be justified if we can show that a normal admissible arc has no set of multipliers with $\lambda_0 = 0$.

Suppose that there were a normal admissible arc with a set of multipliers having $\lambda_0 = 0$. Its function F would have the form

$$F = \lambda_1 \phi_1 + \cdots \lambda_m \phi_m$$

and every set of admissible variations along it would satisfy the equation

$$0 = \int_{x_1}^{x_2} \lambda_a \Phi_a dx = \int_{x_1}^{x_2} (F_{y_i} \eta_i + F_{y_i'} \eta_i') dx = F_{y_i'}(x_2) \eta_i(x_2) - F_{y_i'}(x_1) \eta_i(x_1)$$

on account of the equations of variations (9) and the equations of the theorem above. Since there is a determinant (38) different from zero it follows that the derivatives $F_{y_i'}$ would all vanish at x_1 and x_2 on our extremal. If we define the variables v_i again by equations (30), or by equations (16) with $\lambda_0 = \lambda_{m+1} = \cdots = \lambda_n = 0$, then in equations (17) the coefficients B_i and the initial values $v_i(x_1) = F_{y_i'}(x_1)$ would all vanish. The only continuous solutions of equations (17) under these circumstances are the functions $v_i(x) \equiv 0$, and equations (16) then imply that the multipliers $\lambda_a(x)$ would all vanish identically, which is not the case. Hence a normal admissible arc can not have a set of multipliers with constant multiplier λ_0 equal to zero.

When an extremal arc has multipliers with $\lambda_0 \neq 0$ the multipliers can evidently all be divided by λ_0 to obtain a set of the form $\lambda_0 = 1$, $\lambda_a(x)$. If there were a second set $\lambda_0 = 1$, $\Lambda_a(x)$ the differences 0, $\Lambda_a - \lambda_a$ would also be a set of multipliers for E_{12} with the constant multiplier zero. We have just seen that this is impossible for a normal extremal unless $\Lambda_a - \lambda_a \equiv 0$, so that the multipliers $\lambda_0 = 1$, $\lambda_a(x)$ of a normal extremal E_{12} are unique.

In every neighborhood of a normal admissible arc E_{12} there are an infinity of admissible arcs with the same end-points 1 *and* 2.

To prove this consider the set of $2n$ admissible variations for E_{12} appearing in the determinant (38) and an additional set $\eta_i(x)$. From the results of Section 3 we know that there is a family of admissible arcs $y_i = Y_i(x, b, b_1, b_2, \cdots, b_{2n})$ containing E_{12} when $b = b_1 = \cdots = b_{2n} = 0$ and having the sets $\eta_i(x)$, $\eta_{is}(x)$ $(s = 1, \cdots, 2n)$ as its variations. The $2n$ equations

$$(39) \qquad Y_i(x_1, b, b_1, \cdots, b_{2n}) = y_{i1}, \quad Y_i(x_2, b, b_1, \cdots, b_{2n}) = y_{i2}$$

have the initial solution $(b, b_1, \cdots, b_{2n}) = (0, 0, \cdots, 0)$ at which the functional determinant of their first members with respect to $b_1, \cdots, b_{2n}$ is the determinant (38) and different from zero. Hence by the usual implicit function theorems these equations have solutions $b_s = B_s(b)$ $(s = 1, \cdots, 2n)$ with initial values $B_s(0) = 0$, and the one parameter family of admissible arcs

$$(40) \qquad y_i = Y_i[x, b, B_1(b), \cdots, B_{2n}(b)] = y_i(x, b)$$

defined by them contains the extremal E_{12} for $b = 0$ and has all its curves passing through the points 1 and 2.

COROLLARY. *If each function $\eta_i(x)$ of a set of admissible variations for a normal admissible arc E_{12} vanishes at x_1 and x_2 then there is a one-parameter family of admissible arcs $y_i = y_i(x, b)$ passing through the points* 1 *and* 2, *containing E_{12} for the parameter value $b = 0$, and having the set $\eta_i(x)$ as its variations along E_{12}.*

Let us suppose that in the construction of the family (40) the set $\eta_i(x)$ of the Corollary has been used. Since these functions all vanish at x_1 and x_2 we find from equations (39), by differentiating with respect to b and setting $b = 0$, that

$$\eta_{is}(x_1) B_s'(0) = 0, \quad \eta_{is}(x_2) B_s'(0) = 0.$$

Since the determinant (38) is different from zero these imply that all the derivatives $B_s'(0)$ vanish. Hence the family (40) has the variations

$$y_{ib}(x, 0) = \eta_i(x) + Y_{ib_s} B_s'(0) = \eta_i(x).$$

We know already that the family contains E_{12} for $b = 0$ and has all of its curves passing through 1 and 2.

8. *Problems with variable end-points.** It happens that a number of important applications of the theory of the Lagrange problem are of a slightly

* See Bliss [16].

different type from that described in Section 1. In order to include them as special cases we must permit variable end-points for the curves of the class in which we are seeking a minimum for I. We shall endeavor to find among the arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2)$$

satisfying the system of equations

$$\phi_\alpha(x, y, y') = 0 \qquad (\alpha = 1, \cdots, m < n)$$

and having end-points satisfying the equations

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p \leqq 2n + 2) \tag{41}$$

one which minimizes the integral I. The number p must not exceed the number $2n + 2$ of end values x_1, y_{i1}, x_2, y_{i2} since otherwise equations (41) would in general have no solutions. The problem of Section 1 is a special case of this one with the system (41) having the special form

$$x_1 - \alpha_1 = y_{i1} - \beta_{i1} = x_2 - \alpha_2 = y_{i2} - \beta_{i2} = 0$$

for which p has exactly the value $2n + 2$.

Suppose now that E_{12} is a minimizing arc for the new problem with end values $(x_1, y_{i1}, x_2, y_{i2})$. We add to the hypotheses (a), (b), (c) of Section 1 the assumption

(d) the functions ψ_μ have continuous derivatives up to and including those of the fourth order near the end-values $(x_1, y_{i1}, x_2, y_{i2})$ of E_{12}, and at these values the $p \times (2n + 2)$-dimensional matrix

$$\| \psi_{\mu x_1} \quad \psi_{\mu y_{i1}} \quad \psi_{\mu x_2} \quad \psi_{\mu y_{i2}} \| \tag{42}$$

has rank p.

The last part of this assumption implies that the equations $\psi_\mu = 0$ are all independent.

It is evident that the arc E_{12} must minimize I in the class of admissible arcs having the same end-values, and we can infer at once that it must have a system of multipliers with which it satisfies the necessary conditions deduced in Section 5. But it is important that we should analyse the situation somewhat more closely. Let

$$y_i = y_i(x, b) \qquad [x_1(b) \leqq x \leqq x_2(b)] \tag{43}$$

be a one-parameter family of admissible arcs containing E_{12} for $b = 0$ whose end-values satisfy the equations

$$\psi_\mu\{x_1(b), y_i[x_1(b), b], x_2(b), y_i[x_2(b), b]\} = 0.$$

If we use the notations $x_{1b}(0) = \xi_1$, $x_{2b}(0) = \xi_2$ the derivatives of these equations with respect to b for $b = 0$ are the system

$$\begin{aligned}\Psi_\mu(\xi, \eta) = (\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1})\xi_1 + \psi_{\mu y_{i1}}\eta_i(x_1) \\ + (\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2})\xi_2 + \psi_{\mu y_{i2}}\eta_i(x_2).\end{aligned} \tag{44}$$

These are the *equations of variation* on E_{12} for the functions ψ_μ. When the family (43) is substituted in the integral I we find for the first variation the formula

$$I_1(\xi, \eta) = \int_{x_1}^{x_2} (f_{y_i}\eta_i + f_{y_i'}\eta_i')\,dx + f(x_2)\xi_2 - f(x_1)\xi_1$$

where $f(x_1)$ and $f(x_2)$ are the values of f at the points 1 and 2 on E_{12}. With the help of the expression (18) we may also write

$$\begin{aligned}\lambda_0 I_1(\xi, \eta) = -\int_{x_1}^{x_2} \lambda_r \zeta_r\,dx - \lambda_0 f(x_1)\xi_1 \\ - c_i\eta_i(x_1) + \lambda_0 f(x_2)\xi_2 + \eta_i(x_2)F_{y_i'}(x_2)\end{aligned} \tag{45}$$

where the constants c_i may be arbitrarily chosen.

A set of admissible variations for the present problem is a set ξ_1, ξ_2, $\eta_i(x)$ in which ξ_1 and ξ_2 are arbitrary constants and the functions $\eta_i(x)$ form a set of admissible variations in the sense of Section 3. For a matrix

$$\left\| \begin{matrix} \xi_{11} & \cdot & \cdot & \cdot & \xi_{1,p+1} \\ \xi_{21} & \cdot & \cdot & \cdot & \xi_{2,p+1} \\ \eta_{11} & \cdot & \cdot & \cdot & \eta_{1,p+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1} & \cdot & \cdot & \cdot & \eta_{n,p+1} \end{matrix} \right\|$$

whose columns are sets of admissible variations there exists a family

$$\begin{aligned} y_i = y_i(x, b_1, \cdots, b_{p+1}) \\ x_1(b_1, \cdots, b_{p+1}) \leqq x \leqq x_2(b_1, \cdots, b_{p+1}) \end{aligned} \tag{46}$$

containing E_{12} for $(b_1, \cdots, b_{p+1}) = (0, \cdots, 0)$ and having the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ $(\sigma = 1, \cdots, p+1)$ as its variations along E_{12} with respect to the parameters b_σ. Such a family is that of the Corollary on page 679 with the functions

$$x_\rho(b_1, \cdots, b_{p+1}) = x_\rho + b_\sigma \xi_{\rho\sigma}, \qquad (\rho = 1, 2)$$

adjoined. When the equations of the family (46) are substituted in the integral I and the functions ψ_μ, these become functions of $b_1, \cdots, b_{p+1}$. The first members of the equations

$$I(b_1, \cdots, b_{p+1}) = I_0 + u,$$
$$\psi_\mu(b_1, \cdots, b_{p+1}) = 0$$

must have their functional determinant equal to zero for $(b_1, \cdots, b_{p+1}) = (0, \cdots, 0)$ by the same argument as that on page 682. This determinant is

$$\begin{vmatrix} I_1(\xi_1, \eta_1) & \cdot \quad \cdot \quad \cdot & I_1(\xi_{p+1}, \eta_{p+1}) \\ \Psi_1(\xi_1, \eta_1) & \cdot \quad \cdot \quad \cdot & \Psi_1(\xi_{p+1}, \eta_{p+1}) \\ \cdot & \cdot \quad \cdot \quad \cdot & \cdot \\ \Psi_p(\xi_1, \eta_1) & \cdot \quad \cdot \quad \cdot & \Psi_p(\xi_{p+1}, \eta_{p+1}) \end{vmatrix} \tag{47}$$

in which only the second subscripts of the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}$ have been indicated. From its vanishing we argue as on page 682 that there exists a set of constants $\lambda_0, d_1, \cdots, d_p$ not all zero such that the equation

$$\lambda_0 I_1(\xi, \eta) + d_\mu \Psi_\mu(\xi, \eta) = 0$$

must hold for every set of admissible variations ξ_1, ξ_2, $\eta_i(x)$. With the help of formulas (44) and (45) this becomes

$$\begin{aligned} -\int_{x_1}^{x_2} \lambda_r \zeta_r dx &+ [-\lambda_0 f(x_1) + d_\mu(\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1})]\, \xi_1 \\ &+ [\lambda_0 f(x_2) + d_\mu(\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2})]\, \xi_2 \\ &+ [-c_i + d_\mu \psi_{\mu y_{i1}}]\, \eta_i(x_1) \\ &+ [F_{y_i'}(x_2) + d_\mu \psi_{\mu y_{i2}}]\, \eta_i(x_2) = 0. \end{aligned}$$

After the arbitrary constants c_i in (45) have been so chosen that the coefficients of the terms in $\eta_i(x_1)$ in the last expression all vanish it follows by an argument like that of page 683 that $\lambda_{\mu+1} \equiv \cdots \equiv \lambda_n \equiv 0$ and that the coefficients of ξ_1, ξ_2, $\eta_i(x_2)$ also vanish. This result is equivalent to saying that all the determinants of order $p+1$ of the matrix

$$\begin{Vmatrix} -\lambda_0 f(x_1) & -F_{y_i'}(x_1) & \lambda_0 f(x_2) & F_{y_i'}(x_2) \\ \psi_{\mu x_1} + \psi_{\mu y_{i1}} y_{i1}' & \psi_{\mu y_{i1}} & \psi_{\mu x_2} + \psi_{\mu y_{i2}} y_{i2}' & \psi_{\mu y_{i2}} \end{Vmatrix}$$

are zero, since the constants c_i are from equations (15) the values $F_{y_i'}(x_1)$, and since the multipliers $1, d_1, \cdots, d_p$ satisfy all the linear equations whose coefficients are columns of the matrix. The rank of the last matrix is unchanged when one column is multiplied by a factor and added to another, and $\lambda_0 f = F$ on the admissible arc E_{12}, so that these results can be formulated as follows:

For every minimizing arc for the problem of Lagrange with variable end-points there exists a set of constants c_i $(i = 1, \cdots, n)$ and a function

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1(x)\phi_1 + \cdots + \lambda_m(x)\phi_m$$

such that the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i$$

are satisfied at every point of E_{12}. *The constant* λ_0 *and the functions* $\lambda_\alpha(x)$ $(\alpha = 1, \cdots, m)$ *are not all identically zero on* x_1x_2 *and are continuous except possibly at values of* x *defining corners of* E_{12}. *Furthermore the end-values of* E_{12} *must be such that all the determinants of order* $p+1$ *of the matrix*

$$(48)\qquad \left\| \begin{array}{cccc} -F(x_1) + y_{i1}'F_{y_i'}(x_1) & -F_{y_i'}(x_1) & F(x_2) - y_{i2}'F_{y_i'}(x_2) & F_{y_i'}(x_2) \\ \psi_{\mu x_1} & \psi_{\mu y_{i1}} & \psi_{\mu x_2} & \psi_{\mu y_{i2}} \end{array} \right\|$$

are zero. These last conditions are the so-called transversality conditions.

It is clear that the multipliers λ_0, $\lambda_\alpha(x)$ can not all vanish identically on x_1x_2. Otherwise the constants $d_1, \cdots, d_p$ would have to satisfy the linear equations whose coefficients are the columns of the matrix (42) which has rank p. The constants $\lambda_0, d_1, \cdots, d_p$ would then all be zero which is not the case.

9. *Normal admissible arcs for problems with variable end-points.* A *normal admissible arc* for the problem of Lagrange with variable end-points is one for which there exist p sets of admissible variations $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ $(\mu = 1, \cdots, p)$ such that the matrix

$$(49)\qquad \left| \begin{array}{ccc} \Psi_1(\xi_1, \eta_1) & \cdots & \Psi_1(\xi_p, \eta_p) \\ \cdot & \cdots & \cdot \\ \Psi_p(\xi_1, \eta_1) & \cdots & \Psi_p(\xi_p, \eta_p) \end{array} \right|$$

is different from zero. In the elements of the matrix only the second subscripts of the sets $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ are indicated.

A necessary and sufficient condition that an admissible arc for the problem of Lagrange with variable end-points be normal is that there exists for it no set of multipliers λ_0, $\lambda_\alpha(x)$ *having* $\lambda_0 = 0$ *with which it satisfies the conditions of the last theorem. For a normal extremal arc satisfying the conditions of the last theorem multipliers in the form* $\lambda_0 = 1$, $\lambda_\alpha(x)$ *always exist and in this form they are unique.*

The proof of Section 8 shows that an admissible arc which is not normal has surely a set of multipliers with $\lambda_0 = 0$, since the linear equations whose coefficients are the columns of the determinant (47) have for such an arc solutions $\lambda_0, d_1, \cdots, d_p$ with $\lambda_0 = 0$.

Suppose now that there were a normal admissible arc satisfying the conditions of the theorem of Section 8 and having $\lambda_0 = 0$. Since the matrix preceding (48) is of rank less than $p+1$ we should then have constants d_μ $(\mu = 1, \cdots, p)$ such that

$$
(50) \qquad \begin{aligned}
F(x_1) &= d_\mu(\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1}),\\
F_{y_i'}(x_1) &= d_\mu \psi_{\mu y_{i1}},\\
-F(x_2) &= d_\mu(\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2}),\\
-F_{y_i'}(x_2) &= d_\mu \psi_{\mu y_{i2}}.
\end{aligned}
$$

The numbers $F(x_1)$, $F(x_2)$ would be zero since $\lambda_0 = 0$ and along an admissible arc $F = \lambda_0 f$. After multiplying these equations respectively by ξ_1, $\eta_i(x_1)$, ξ_2, $\eta_i(x_2)$ and adding we should have

$$
\eta_i(x_1) F_{y_i'}(x_1) - \eta_i(x_2) F_{y_i'}(x_2) = d_\mu \Psi_\mu(\xi, \eta).
$$

The first member of this equation would vanish for every set of admissible variations $\eta_i(x)$, as was proved in Section 7, page 688, and the second member would necessarily have the same property. Since there is a determinant (49) different from zero we should then have $d_\mu = 0$ for every μ, and equations (50) show that $F_{y_i'}(x_1)$ and $F_{y_i'}(x_2)$ would all vanish. As in Section 7, page 688, this would necessitate the vanishing of λ_0, $\lambda_a(x)$ which is impossible. The proof of the uniqueness of the multipliers $\lambda_0 = 1$, $\lambda_a(x)$ is precisely that of Section 7.

In every neighborhood of a normal admissible arc E_{12} for the Lagrange problem with variable end-points there is an infinity of admissible arcs satisfying the end conditions $\psi_\mu = 0$.

The proof is similar to that of the corresponding theorem in Section 7. Select arbitrarily an admissible set of variations ξ_1, ξ_2, $\eta_i(x)$ and p other such sets $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ with determinant (49) different from zero. There is a $p+1$-parameter family

$$
(51) \qquad \begin{gathered}
y_i = Y_i(x, b, b_1, \cdots, b_p)\\
X_1(b, b_1, \cdots b_p) \leqq x \leqq X_2(b, b_1, \cdots, b_p)
\end{gathered}
$$

of admissible arcs containing E_{12} for $(b, b_1, \cdots, b_p) = (0, 0, \cdots, 0)$ and having the sets ξ_1, ξ_2, $\eta_i(x)$ and $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ as its variations along E_{12}. The existence of the functions Y_i is a consequence of the corollary of Section 3 above, and we may take $X_\rho = x_\rho + b\xi_\rho + b_\mu \xi_{\rho\mu}$ $(\rho = 1, 2)$. Each function ψ_μ becomes a function $\psi_\mu(b, b_1, \cdots, b_p)$ when the functions (51) defining these arcs are substituted. The equations

$$
(52) \qquad \psi_\mu(b, b_1, \cdots, b_p) = 0
$$

have the initial solution $(b, b_1, \cdots b_\mu) = (0, 0, \cdots, 0)$ at which the functional determinant of their first members with respect to $b_1, \cdots, b_\mu$ is the determinant (49) different from zero. Hence these equations have p solutions $b_\mu = B_\mu(b)$ with initial values $B_\mu(0) = 0$. The one-parameter family

$$(53) \qquad y_i = Y_i[x, b, B_1(b), \cdots, B_p(b)] = y_i(x, b)$$
$$x_1(b) \leqq x \leqq x_2(b)$$

where

$$x_\rho(b) = X_\rho[b, B_1(b), \cdots, B_p(b)] \qquad (\rho = 1, 2)$$

contains E_{12} for $b = 0$ and satisfies the equations $\psi_\mu = 0$.

COROLLARY. *If a set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ for a normal admissible arc E_{12} for the Lagrange problem with variable end-points satisfies the equations $\Psi_\mu(\xi, \eta) = 0$, then there exists a one parameter family*

$$y_i = y_i(x, b), \quad x_1(b) \leq x \leq x_2(b)$$

of admissible arcs satisfying the end-conditions $\psi_\mu = 0$, containing E_{12} for the parameter value $b = 0$, and having the set ξ_1, ξ_2, $\eta_i(x)$ as its variations along E_{12}.

If the set ξ_1, ξ_2, $\eta_i(x)$ of the Corollary is used in the construction of the family (53) then we find, by differentiating equations (52) with respect to b and setting $b = 0$, that

$$\Psi_\mu(\xi, \eta) + \Psi_\mu(\xi_\nu, \eta_\nu) B_\nu'(0) = 0.$$

But since the first terms in these equations vanish, and since the determinant (49) is different from zero, it follows that $B_\mu'(0) = 0$ for every μ. Hence the variations of the family (53) are the functions

$$y_{ib}(x, 0) = \eta_i(x) + Y_{ib_\mu} B_\mu'(0) = \eta_i(x),$$
$$x_{\rho b}(0) = \xi_\rho + X_{\rho b_\mu} B_\mu'(0) = \xi_\rho, \qquad (\rho = 1, 2)$$

as required in the Corollary.

CHAPTER II.

APPLICATIONS OF THE EULER-LAGRANGE MULTIPLIER RULE.

10. *The brachistochrone in a resisting medium.* Analytically the problem of the brachistochrone in a plane and in a resisting medium is, as we have seen in Section 2, that of finding among the arcs

$$y = y(x), \quad v = v(x) \qquad (x_1 \leq x \leq x_2)$$

satisfying the conditions

$$
(54)\qquad \begin{gathered} vv' - gy' + R(v)(1+y'^2)^{1/2} = 0, \\ x_1 - \alpha_1 = y_1 - \beta_1 = v_1 - \gamma = x_2 - \alpha_2 = y_2 - \beta_2 = 0, \end{gathered}
$$

one which minimizes the integral

$$
I = \int_{x_1}^{x_2} (1/v)(1+y'^2)^{1/2} dx.
$$

In these expressions primes denote derivatives with respect to x. To apply the Euler-Lagrange rule and the transversality conditions of Section 8 we construct the function

$$
\begin{aligned} F &= (1/v)(1+y'^2)^{1/2} + \lambda[vv' - gy' + R(1+y'^2)^{1/2}] \\ &= H(1+y'^2)^{1/2} + \lambda(vv' - gy') \end{aligned}
$$

where H is a convenient symbol * for the expression

$$
(55)\qquad H = (1/v) + \lambda R(v).
$$

The differential equations of the normal extremals are then easily found to be

$$
(56)\quad H(dy/ds) = \lambda g + a, \quad v(d\lambda/ds) = H_v, \quad v(dv/ds) = g(dy/ds) - R
$$

where s is the length of arc defined by the equation

$$
ds = (1+y'^2)^{1/2} dx
$$

and a is a new constant of integration. By eliminating dy and ds from equations (56) we find

$$
H(H_v dv + R d\lambda) = (g\lambda + a) g d\lambda,
$$

which gives at once, since $H_\lambda = R$, the relation

$$
(57)\qquad H^2 = (g\lambda + a)^2 + b^2
$$

where b is a second constant of integration. The constant can be taken squared since the first equation (56) shows that H^2 is always greater than $(\lambda g + a)^2$.

Equations (56) and (57) give further

$$
(58)\qquad \begin{aligned} \frac{dy}{dv} &= \frac{dy}{ds}\,\frac{ds}{dv} = \frac{v(\lambda g + a)}{g(\lambda g + a) - RH}, \\ \frac{dx}{dv} &= \frac{dx}{ds}\,\frac{ds}{dv} = \left[1 + \left(\frac{dy}{ds}\right)^2\right]^{1/2} \frac{ds}{dv} = \frac{bv}{g(\lambda g + a) - RH}. \end{aligned}
$$

* Bolza [3, p. 577].

Equation (57) is quadratic in λ and when its solution $\lambda = \lambda(v, a, b)$ is substituted in the last equations the values of x and y may be found by quadratures in the form

$$x = \phi(v, a, b) + c, \quad y = \psi(v, a, b) + d, \tag{59}$$

where c and d are again constants of integration. These are the equations of the minimizing arc in parametric form.

It is very easy to set up the matrix (48) for our function F and the five end conditions. It is a square matrix with six rows and columns and its vanishing prescribes the single condition $\lambda(x_2)v(x_2) = 0$. From the equation (57) multiplied by v^2 and equation (55) we then find at $x = x_2$ that $v_2{}^2(a^2 + b^2) = 1$. For the determination of v_2 and the four constants of integration in equations (59) we have therefore in accordance with conditions (54) the five equations

$$\begin{gathered} \phi(v_1, a, b) + c = \alpha_1, \qquad \phi(v_2, a, b) + c = \alpha_2, \\ \psi(v_1, a, b) + d = \beta_1, \qquad \psi(v_2, a, b) + d = \beta_2, \\ v_2{}^2(a^2 + b^2) = 1. \end{gathered} \tag{60}$$

If the resistance function $R(v)$ were known we should now have in equations (57), (56), and (60) the mathematical mechanism for determining possible normal minimizing curves. The adjective possible is used here because the conditions deduced so far have only been shown to be necessary for a normal minimizing arc. They have not been proved to be sufficient to insure a minimum.

11. *Parametric problems in space.* Let us now consider space curves whose equations are given in the parametric form

$$x = x(s), \quad y = y(s), \quad z = z(s) \qquad (s_1 \leq s \leq s_2). \tag{61}$$

The problem to be studied is that of finding among the arcs of this type which satisfy the equation

$$x'^2 + y'^2 + z'^2 - 1 = 0 \tag{62}$$

and join two given points 1 and 2 in xyz-space, one which minimizes an integral of the form

$$I = \int_{s_1}^{s_2} f(x, y, z, x', y', z')\, ds.$$

Primes now denote differentiation with respect to s. Equation (62) restricts the parameter s to be the length of arc measured along the curve (61). If

we agree to measure this length always from the point 1 then the conditions for the curve (61) to pass through 1 and 2 are

$$s_1 = x_1 - \alpha_1 = y_1 - \beta_1 = z_1 - \gamma_1 = x_2 - \alpha_2 = y_2 - \beta_2 = z_2 - \gamma_2 = 0$$

where $(\alpha_1, \beta_1, \gamma_1)$ and $(\alpha_2, \beta_2, \gamma_2)$ are the coördinates of these points. Evidently our problem is one with a variable end-point in $sxyz$-space since s_2 is undetermined.

The function F for normal minimizing arcs is

$$F = f + (\lambda/2)(x'^2 + y'^2 + z'^2 - 1)$$

and the differential equations determining such arcs are

$$\begin{aligned} &f_x - (d/ds) f_{x'} - \lambda' x' - \lambda x'' = 0, \\ &f_y - (d/ds) f_{y'} - \lambda' y' - \lambda y'' = 0, \\ &f_z - (d/ds) f_{z'} - \lambda' z' - \lambda z'' = 0, \\ &x'^2 + y'^2 + z'^2 = 1. \end{aligned} \tag{63}$$

The sum of the first three of these multiplied, respectively, by x', y', z' gives, with the help of the last one,

$$(d/ds)(f - x' f_{x'} - y' f_{y'} - z' f_{z'} - \lambda) = 0. \tag{64}$$

The matrix (48) for this problem has eight rows and columns and the vanishing of its determinant demands that at the value s_2

$$\lambda = f - x' f_{x'} - y' f_{y'} - z' f_{z'}. \tag{65}$$

On account of equation (64) this must be an identity in s.

A very important case is the one for which the function f is positively homogeneous and of the first order in x', y', z', i. e. the one for which the equation

$$f(x, y, z. kx', ky', kz') = kf(x, y, z, x', y', z') \tag{66}$$

is an identity in its arguments for all $k > 0$. The integral I then has the same value for all parametric representations of the arc (61). The integrands of the length integral and of many other integrals important in the applications of the theory of the Lagrange problem satisfy this condition. When equation (66) is differentiated for k, and the substitution $k = 1$ afterward made, we find the identity

$$x' f_{x'} + y' f_{y'} + z' f_{z'} = f. \tag{67}$$

From equation (65) it is evident that in this case $\lambda = 0$ and equations (63) become

$$f_x - (d/ds) f_{x'} = 0, \quad f_y - (d/ds) f_{y'} = 0, \quad f_z - (d/ds) f_{z'} = 0, \tag{68}$$

$$x'^2 + y'^2 + z'^2 - 1 = 0. \tag{69}$$

Only three of these can be independent, since one finds readily that

$$x'P + y'Q + z'R = (d/ds)(f - x'f_{x'} - y'f_{y'} - z'f_{z'}) = 0$$

where P, Q, R are symbols for the first members of equations (68).

12. *Isoperimetric problems.* Suppose that we seek to find in the class of arcs

$$y = y(x) \qquad (x_1 \leqq x \leqq x_2)$$

joining two given points and satisfying relations of the form

$$\int_{x_1}^{x_2} g_i(x, y, y')\,dx = l_i \qquad (i = 1, \cdots, n) \tag{70}$$

one which minimizes an integral

$$I = \int_{x_1}^{x_2} f(x, y, y')\,dx.$$

We can transform such a problem into a Lagrange problem by introducing new variables

$$z_i(x) = \int_{x_1}^{x} g_i(x, y, y')\,dx. \tag{71}$$

The problem just stated is then equivalent to that of finding in the class of arcs

$$y = y(x), \quad z_i = z_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2)$$

satisfying the conditions

$$\begin{aligned} &g_i(x, y, y') - z_i' = 0, \\ &y(x_1) = y_1, \quad y(x_2) = y_2, \\ &z_i(x_1) = 0, \quad z_i(x_2) = l_i, \qquad (i = 1, \cdots, n) \end{aligned} \tag{72}$$

one which minimizes I.

The function F for a normal minimizing arc for this problem has the form

$$F = f + \lambda_i (g_i - z_i') \tag{73}$$

and the differential equations determining such an arc are

$$F_y - (d/dx) F_{y'} = 0 \tag{74}$$

and the n equations

$$F_{z_i} - (d/dx)F_{z_i'} = (d\lambda_i/dx) = 0$$

which show that the multipliers λ_i are in this case all constants. The solutions of equations (74) form a family of the type

$$y = y(x, a, b, \lambda_1, \cdots, \lambda_n).$$

It contains $n+2$ arbitrary constants, and that is precisely the number of relations which the end-conditions (72) impose upon them as one readily verifies. It is evident that the equation (74) is unaltered if we think of the function F in it as defined by the equation

$$F = f + \lambda_i g_i \tag{75}$$

instead of equation (73).

For a minimizing arc which is not normal there would be a function F defined by equation (75) without the first term. It is clear that the equation (74) would then be defining the minimizing arcs for the problem of minimizing one of the integrals (70), say the first one, in the class of curves joining 1 with 2 and keeping the others constant. An arc E_{12} satisfying equations (74) and these conditions would in general be a minimizing arc for this problem, and it is evident that in that case there could be no other arc near E_{12} giving the first integral its minimum value l_1. Hence in a neighborhood of E_{12} the class of arcs joining 1 with 2 and satisfying conditions (70) would consist of E_{12} alone, and the original minimum problem would be a very trivial one in that neighborhood. Evidently the normal minimizing arcs are by far the most important ones. A similar but somewhat more complicated argument justifies the definition of normal minimizing arcs for the general Lagrange problem given in the preceding sections.

13. *The hanging chain.* It is a principle of mechanics that a chain suspended on two pegs will hang so that its center of gravity is as low as possible. In Section 2 it was seen that the form of the chain is therefore that of a minimizing arc for the problem in which we seek among the arcs $y = y(x)$ $(x_1 \leq x \leq x_2)$ satisfying the conditions

$$y(x_1) = y_1, \quad y(x_2) = y_2, \quad \int_{x_1}^{x_2} (1 + y'^2)^{1/2} dx = l, \tag{76}$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2} dx.$$

The function F for a minimizing arc has the form

$$F = (y + \lambda)(1 + y'^2)^{1/2}$$

and since λ is now constant the differential equation (74) is equivalent to

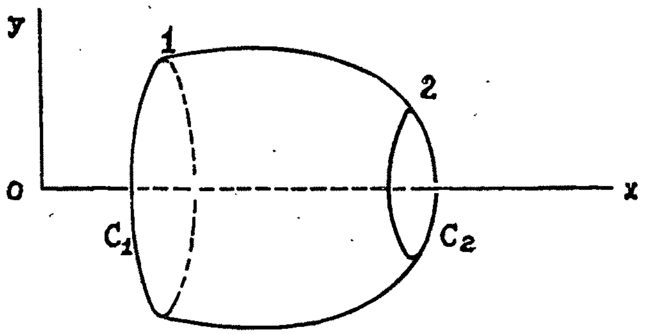
$$F - y'F_{y'} = (y + \lambda)/(1 + y'^2)^{1/2} = b.$$

The integration of this equation has been many times discussed * and its solutions are the catenaries

$$y + \lambda = b\, ch[(x - a)/b].$$

This is a larger family than that of the catenaries for the problem of finding a minimum surface of revolution since it contains an arbitrary constant λ besides a and b. The extra constant is needed, however, for the problem of the hanging chain since there are three conditions (76) to be satisfied for that problem instead of the first two only.

14. *Soap films enclosing a given volume.* Let C_1 and C_2 be two circular discs with a common axis whose edges are joined by a soap film. It is well known that when the volume of air inclosed by the discs and the film is a



fixed constant k the form of the film surface will be that of a surface of revolution enclosing the volume k and having a minimum surface area. To determine the shape of the film we must seek therefore among the arcs $y = y(x)$ $(x_1 \leq x \leq x_2)$ satisfying the conditions

$$y(x_1) = y_1, \quad y(x_2) = y_2, \quad \int_{x_1}^{x_2} y^2 dx = k/\pi$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2} dx.$$

* See, for example, Bliss [5, p. 91].

The function F is $F=y(1+y'^2)^{1/2}+\lambda y^2$ and the equation (74) is equivalent to

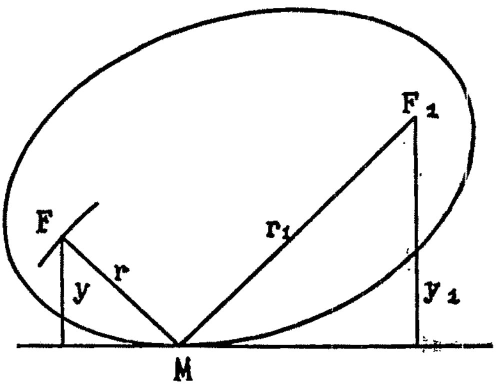
$$F-y'F_{y'}=y/(1+y')^{1/2}-\lambda y^2=c. \tag{77}$$

If we solve this equation for y' and separate the variables we find the solution in the form

$$x=\int\{(c-\lambda y^2)/[y^2-(c-\lambda y^2)^2]^{1/2}\}dy+d.$$

The integral here is an elliptic integral which can be treated by well known methods.

The solutions of equations (77) can be characterized geometrically in an interesting fashion.* If an ellipse rolls on a straight line, as in the accompanying figure, its focus F describes a curve whose tangent is at every point perpendicular to FM. The coördinates (x,y) of F, and (x_1,y_1) of F_1, therefore satisfy the equations



$$y=r(dx/ds),\quad y_1=r_1(dx/ds)$$

since by a well known property of the ellipse the angles made by r and r_1 with the tangent at M are equal. The equations

$$r+r_1=2a,\quad yy_1=b^2$$

express two further well known properties of an ellipse, and elimination of r, r_1, y_1 from these and the preceding ones gives the differential equation

$$y^2-2ay(dx/ds)+b^2=0$$

* See, for example, Moigno-Lindelöf [6, p. 220].

for the locus of the point F. Equation (77) is identical with this if we set $\lambda = -1/2a$, $c = b^2/2a$. It can similarly be shown that for suitable determinations of λ and c equation (77) is also satisfied by the locus of the focus of a parabola or a hyperbola which rolls on the x-axis. The curves generated as described above by the foci of conics rolling on the x-axis are called unduloids and nodoids.

15. *The case when the functions ϕ_α contain no derivatives.* The problem of this section is that of finding among the arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2) \tag{78}$$

joining the two given points 1 and 2 and satisfying a set of equations of the form

$$\phi_\alpha(x, y_1, \cdots, y_n) = 0 \qquad (\alpha = 1, \cdots, m < n)$$

one which minimizes an integral

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx.$$

Let E_{12} be a particular arc whose minimizing properties are to be studied. It is always presupposed that in a neighborhood of the set of elements (x, y, y') on E_{12} the functions f, ϕ_α have continuous partial derivatives, say of the first four orders, and that the matrix $\| \partial\phi_\alpha / \partial y_i \|$ has rank m at every point of E_{12}.

In order to give this problem the usual Lagrange form we replace it by an equivalent one as follows. We may suppose without loss of generality that at the point 2 the determinant $| \partial\phi_\alpha / \partial y_\beta |$ is one of those of the matrix $\| \partial\phi_\alpha / \partial y_i \|$ which is different from zero. Then we seek to find among the arcs (78) satisfying the conditions

$$d\phi_\alpha / dx = \phi_{\alpha x} + \phi_{\alpha y_i} y_i' = 0, \tag{79}$$

$$\begin{gathered} x_1 - \alpha_1 = y_{i1} - \beta_{i1} = x_2 - \alpha_2 = y_{r2} - \beta_{r2} = 0 \\ (i = 1, \cdots, n;\ r = m + 1, \cdots, n) \end{gathered} \tag{80}$$

one which minimizes I. The coördinates (α_1, β_{i1}) and (α_2, β_{i2}) are those of the points 1 and 2 and necessarily satisfy the equations $\phi_\alpha = 0$. The new problem is evidently equivalent to the old one, at least in a neighborhood of E_{12}, since every arc (78) which joins 1 with 2 and satisfies the equations $\phi_\alpha = 0$ also satisfies (79) and (80); and since, conversely, every arc sufficiently near E_{12} and satisfying (79) and (80) will also satisfy the equations

$\phi_a = 0$ and pass through 1 and 2. This follows because the last $n - m + 1$ equations (80) and the equations $\phi_a = 0$ at 2 imply $y_{a2} - \beta_{a2} = 0$.

Every extremal arc for the new problem is necessarily normal. The determinant analogous to (49) for the end-conditions (80) is in fact

$$\begin{vmatrix} \xi_{11} & \cdot\ \cdot\ \cdot & \xi_{1p} \\ \eta_{11}(x_1) & \cdot\ \cdot\ \cdot & \eta_{1p}(x_1) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \eta_{n1}(x_1) & \cdot\ \cdot\ \cdot & \eta_{np}(x_1) \\ \xi_{21} & \cdot\ \cdot\ \cdot & \xi_{2p} \\ \eta_{m+1,1}(x_2) & \cdot\ \cdot\ \cdot & \eta_{m+1,p}(x_2) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \eta_{n1}(x_2) & \cdot\ \cdot\ \cdot & \eta_{np}(x_2) \end{vmatrix}$$

where $p = 2n - m + 2$, and we can prove that the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ $(\sigma = 1, \cdot \cdot \cdot, p)$ can be chosen so that this determinant is different from zero. The equations of variation are in fact readily seen to be the equations

$$(d/dx)\phi_{ay_i}\eta_i = 0$$

which are equivalent to the system

$$\phi_{ay_i}(x)\eta_i(x) = \phi_{ay_i}(x_1)\eta_i(x_1). \tag{81}$$

If the end-values $\eta_i(x_1)$, $\eta_r(x_2)$ are selected arbitrarily these equations determine uniquely the end-values $\eta_a(x_2)$ since the determinant $|\,\partial\phi_a/\partial y_\beta\,|$ is by hypothesis different from zero at the point 2. Then the equations (81) and

$$\phi_{ry_i}(x)\eta_i(x) = \zeta_r(x), \tag{82}$$

where the auxiliary functions $\phi_r(x, y)$ are chosen so that the determinant $|\,\partial\phi_i/\partial y_k\,|$ is different from zero along E_{12}, determine the end-values $\zeta_r(x_1)$, $\zeta_r(x_2)$ uniquely when $\eta_i(x_1)$, $\eta_r(x_2)$ are given. If functions $\zeta_r(x)$ are chosen with the end-values $\zeta_r(x_1)$, $\zeta_r(x_2)$ but otherwise arbitrarily then equations (81) and (82) determine uniquely a corresponding set of variations $\eta_i(x)$ with the arbitrarily prescribed end-values $\eta_i(x_1)$, $\eta_r(x_2)$. Since ξ_1 and ξ_2 are arbitrary it is evident that the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ can be chosen so that the determinant above is different from zero.

The function F for the Euler-Lagrange multiplier rule of the new problem can be taken in the form

$$F = f + \mu_a(\phi_{ax} + \phi_{ay_k}y_k').$$

By a simple calculation the Euler-Lagrange equations are found to be

$$f_{y_i} - (d/dx)f_{y_i'} - \mu_a'\phi_{ay_i} = 0.$$

If we set $\lambda_a = -\mu_a'$ these are equivalent to the Euler-Lagrange equations calculated for the function

$$F = f + \lambda_a \phi_a$$

and we have the following result:

For the problem of finding among the arcs $y_i = y_i(x)$ $(i = 1, \cdots, n;$ $x_1 \leq x \leq x_2)$ *joining two given points and satisfying the equations*

$$\phi_a(x, y) = 0 \qquad (\alpha = 1, \cdots, m < n)$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx,$$

the extremal arcs all satisfy $n + m$ *equations of the form*

$$F_{y_i} - (d/dx) F_{y_i'} = 0, \quad \phi_a = 0$$

where F *is a function of the form* $F = f + \lambda_a \phi_a$.

16. *Geodesics on a surface.** The problem of finding the shortest curve joining two given points on a surface is analytically that of finding among the arcs

$$x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t_1 \leq t \leq t_2)$$

satisfying the equation

$$\phi(x, y, z) = 0 \tag{83}$$

of the surface and joining the two given points, one which minimizes the integral

$$I = \int_{t_1}^{t_2} (x'^2 + y'^2 + z'^2)^{1/2} dt.$$

The function F for this problem, according to the results of the last section, is

$$F = (x'^2 + y'^2 + z'^2)^{1/2} + \lambda\phi$$

and the Euler-Lagrange equations are $\phi = 0$ and

$$\begin{aligned}
(d/dt) F_{x'} - F_x &= d/dt[x'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda\phi_x = 0,\\
(d/dt) F_{y'} - F_y &= d/dt[y'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda\phi_y = 0,\\
(d/dt) F_{z'} - F_z &= d/dt[z'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda\phi_z = 0.
\end{aligned}$$

If these are written in the form

* Bolza [3, p. 553].

$$d^2x/ds^2 = \mu\phi_x, \quad d^2y/ds^2 = \mu\phi_y, \quad d^2z/ds^2 = \mu\phi_z, \quad \phi = 0,$$

where s is the length of arc, they express the fact that at each point of a minimizing arc the principal normal of the arc must coincide with the normal to the surface. Curves which have this property are called *geodesic lines* on the surface. Shortest arcs on a surface must always be sought among the geodesics.

For a sphere the equation (83) has the form

$$x^2 + y^2 + z^2 - 1 = 0$$

and the further equations of the geodesics are

$$d^2x/ds^2 = \mu x, \quad d^2y/ds^2 = \mu y, \quad d^2z/ds^2 = \mu z. \tag{84}$$

Let us determine constants a, b, c so that the expression

$$u = ax + by + cz$$

vanishes with its first derivative at one point of a geodesic on the sphere. Then u must be identically zero on the geodesic since the equation $u_{ss} = \mu u$ is a consequence of equations (84), and since the only solution of this last equation which can vanish with its derivative is $u \equiv 0$. It follows readily that the geodesics on a sphere are great circles cut out of the sphere by the planes $u = 0$.

17. *Brachistochrone on a surface.** Consider a particle of mass m moving in a field of force of such nature that when the particle is at the point (x, y, z) the force acting on it has the projections

$$mX = m(\partial U/\partial x), \quad mY = m(\partial U/\partial y), \quad mZ = m(\partial U/\partial z) \tag{85}$$

on the three coördinate axes, where U is a function of the coördinates x, y, z only. A constant gravitational field in the direction of the negative z-axis, for example, would have

$$X = 0, \quad Y = 0, \quad Z = -g, \quad U = -gz.$$

If a particle were constrained to move on a curve in such a field we should have the force in the direction of the tangent expressed in the two forms

$$mv' = m\,[X(dx/ds) + Y(dy/ds) + Z(dz/ds)]$$

where v is the velocity in the tangent direction, s is the length of arc measured along the curve, and the prime denotes a derivative with respect to the time t. Since $v = ds/dt$ this gives

* Moigno-Lindelöf [6, p. 301].

$$
(86) \qquad \begin{aligned} vv' &= Xx' + Yy' + Zz' = U', \\ v^2 &= 2U + c = 2(U - U_1) + v_1^2, \end{aligned}
$$

where U_1 and v_1 are values of U and v at an initial point 1. For a particle started at 1 with the velocity v_1 the velocity v at a point (x, y, z) is evidently a function of x, y, z and the same for all arcs joining 1 with this point. For an arc

$$
(87) \qquad x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t_1 \leq t \leq t_2)
$$

joining two fixed points 1 and 2 the time of descent of a particle starting at 1 with the velocity v_1 is

$$
T = \int_{s_1}^{s_2} ds/v = \int_{t_1}^{t_2} (1/v)(x'^2 + y'^2 + z'^2)^{1/2} dt
$$

where v is the function of x, y, z defined in equation (86).

The problem of finding an arc of quickest descent from a point 1 to a point 2 on a surface

$$
(88) \qquad \phi(x, y, z) = 0
$$

for a particle starting at 1 with a given velocity v_1 is equivalent analytically to that of finding among the arcs (87) joining the two given points and satisfying the equation (88), one which minimizes the integral T.

The function F for this problem is

$$
F = (1/v)(x'^2 + y'^2 + z'^2)^{1/2} + \lambda\phi
$$

and the Euler-Lagrange equations have the form

$$
\begin{aligned}
\frac{d}{dt} F_{x'} - F_x &= \frac{d}{dt}\,\frac{1}{v}\,\frac{dx}{ds} + \frac{v_x}{v^2}\,\frac{ds}{dt} - \lambda\phi_x = 0, \\
\frac{d}{dt} F_{y'} - F_y &= \frac{d}{dt}\,\frac{1}{v}\,\frac{dy}{ds} + \frac{v_y}{v^2}\,\frac{ds}{dt} - \lambda\phi_y = 0, \\
\frac{d}{dt} F_{z'} - F_z &= \frac{d}{dt}\,\frac{1}{v}\,\frac{dz}{ds} + \frac{v_z}{v^2}\,\frac{ds}{dt} - \lambda\phi_z = 0
\end{aligned}
$$

to which must be adjoined the equation $\phi = 0$. When multiplied through by dt/ds the equations above become

$$
\begin{aligned}
-(v_s/v^2)x_s + (1/v)x_{ss} + (v_x/v^2) - \mu\phi_x &= 0, \\
-(v_s/v^2)y_s + (1/v)y_{ss} + (v_y/v^2) - \mu\phi_y &= 0, \\
-(v_s/v^2)z_s + (1/x)z_{ss} + (v_z/v^2) - \mu\phi_z &= 0.
\end{aligned}
$$

Multiplied respectively by the direction cosines l, m, n of the direction tan-

gent to the surface, perpendicular to the extremal, and making an acute angle with its principal normal, these give

$$(1/v)(lx_{ss} + my_{ss} + nz_{ss}) + (1/v^2)(v_x l + v_y m + v_z n) = 0$$

from which we can show that

$$(v^2/\rho)\cos\alpha + R\cos\beta = 0 \tag{89}$$

where ρ is the radius of curvature of the curve, α the angle between the radius and the direction $l:m:n$, R the total impressed force, and β the angle between the force and $l:m:n$. This result follows immediately since the numbers ρx_{ss}, ρy_{ss}, ρz_{ss} are the three direction cosines of the principal normal to the curve on which the radius ρ lies, and since from equations (86)

$$vv_x = U_x, \quad vv_y = U_y, \quad vv_z = U_z$$

and U_x, U_y, U_z are the projections on the coördinate axes of the force R. The equation (89) justifies the following characteristic property of brachistochrones on a surface:

Consider a surface $\phi(x, y, z) = 0$ lying in a field of force whose vector at (x, y, z) has magnitude R and components X, Y, Z defined by a force function $U(x, y, z)$, as indicated in equations (85). *The centrifugal force of a particle moving on a curve is by definition directed in the direction opposite to that of the radius ρ of the first curvature, and has magnitude v^2/ρ where v is the velocity of the particle. Equation* (89) *shows that at each point of a brachistochrone curve on the surface $\phi = 0$ the projection of the centrifugal force on the particular normal to the curve which is also tangent to the surface, is equal to the projection on that same line of the impressed force R.*

This is a characteristic property of brachistochrones. Equation (89) shows that the radius of geodesic curvature $\rho_g = \rho\sec\alpha$ is defined by the equation

$$1/\rho_g = -(R/v^2)\cos\beta. \tag{90}$$

On a surface whose equations are in parametric form with parameters u, v the geodesic curvature of an arc defined by an equation $v = v(u)$ is expressed in terms of $v(u)$, $v'(u)$, $v''(u)$ while the quantities in the second members of the last equation involve only $v(u)$ and $v'(u)$. This equation is consequently a differential equation of the second order. Through each point and direction on the surface there passes therefore one and only one extremal arc for the brachistochrone problem. One can readily verify that the equation

(90) is satisfied by the brachistochrones on a plane which are the well-known cycloids.

18. *The curve of equilibrium of a chain hanging on a surface.** Let us accept from the theories of mechanics the statement that the potential energy of a chain of the form

$$(91) \qquad x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t_1 \leq t \leq t_2)$$

in a field of force like the one described in the last section is

$$P = -\int_{s_1}^{s_2} U ds = -\int_{t_1}^{t_2} U(x'^2 + y'^2 + z'^2)^{1/2} dt,$$

and the statement that a chain at rest will be in equilibrium when the potential energy is a minimum. The problem of finding the position of equilibrium of a chain of given length l joining two given points 1 and 2 and lying on a surface

$$\phi(x, y, z) = 0$$

in such a field is then that of finding among the arcs (91) joining 1 with 2 and satisfying the conditions

$$\int_{t_1}^{t_2} (x'^2 + y'^2 + z'^2)^{1/2} dt = l, \quad \phi(x, y, z) = 0$$

one which minimizes the integral P. In a gravitational field the value of U is $-gz$.

This problem is partly of the isoperimetric and partly of the Lagrange type. By methods used above one readily verifies that its function F now has the form

$$F = (U + \lambda)(x'^2 + y'^2 + z'^2)^{1/2} + \mu\phi,$$

where λ is a constant, and that its extremal arcs satisfy $\phi = 0$ and the equations

$$\begin{aligned}
d/dt[(U + \lambda)x'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_x(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi &= 0,\\
d/dt[(U + \lambda)y'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_y(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi_y &= 0,\\
d/dt[(U + \lambda)z'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_z(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi_z &= 0.
\end{aligned}$$

These are equivalent to

$$\begin{aligned}
U_s x_s + (U + \lambda)x_{ss} - U_x - \nu\phi_x &= 0,\\
U_s y_s + (U + \lambda)y_{ss} - U_y - \nu\phi_y &= 0,\\
U_s z_s + (U + \lambda)z_{ss} - U_z - \nu\phi_z &= 0.
\end{aligned}$$

* Moigno-Lindelöf [6, p. 313].

Multiplied respectively by the direction cosines l, m, n of the direction tangent to the surface, perpendicular to the extremal, and making an acute angle with its principal normal, these give

$$(U + \lambda) \cos \alpha / \rho = R \cos \beta,$$

or

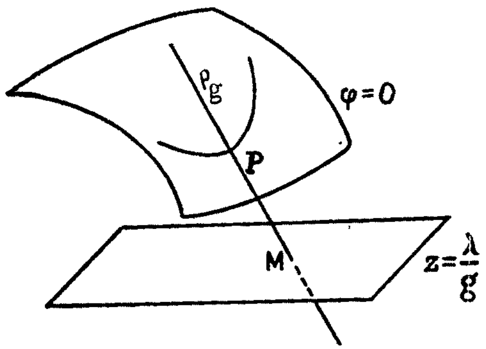
$$\rho_g = (U + \lambda) \sec \beta / R,$$

where ρ, ρ_g, α, β have the significance of the last section. Like the equation (90) this defines a two-parameter family of extremals arcs on the surface $\phi = 0$.

For the particular case of a gravitational field of force $U = -gz$, $R = g$, and β is the angle between the negative z-axis and the direction $l:m:n$ so that $\cos \beta = -n$. Hence in this case

$$\rho_g = [(z - \lambda)/g]/n$$

which says that at each point of a curve of equilibrium the radius of geodesic curvature is equal to the segment PM in the figure, bounded on the line



$l:m:n$ perpendicular to the curve and tangent to the surface $\phi = 0$ by the point P and the plane $z = \lambda/g$. This is a well known property of a catenary $y = c + b\ ch\ [(x - a)/b]$, which is the curve of a hanging chain in a vertical plane. The surface $\phi = 0$ is in this case the xy-plane, the radius ρ_g is the radius of curvature of the catenary, and the plane $z = \lambda/g$ is to be represented by the line $y = c$. The radius of curvature at a point P of the catenary is equal to the intercept on the normal to the catenary at P between the point P and the line $y = c$.

19. *Hamilton's principle.** Suppose that the n particles whose coördinates and masses are x_i, y_i, z_i, m_i $(i = 1, \cdots, n)$ move in a field of force

* Bolza [3, p. 554].

in space such that the force acting at any instant on the i-th particle has components

$$X_i = U_{x_i}, \quad Y_i = U_{y_i}, \quad Z_i = U_{z_i},$$

where U is a function of the time t and the $3n$ coördinates x_i, y_i, z_i. Suppose further that the motions of the particles are restricted by conditions of the form

$$\phi_\alpha = 0 \qquad (\alpha = 1, \cdots, m < 3n),$$

where the functions ϕ_α also depend upon t and the coördinates. The differential equations of motion of the particles, as established in treatises in mechanics, are

$$\begin{aligned} m_i x_i'' &= U_{x_i} + \Sigma_\alpha \lambda_\alpha \phi_{\alpha x_i}, \\ m_i y_i'' &= U_{y_i} + \Sigma_\alpha \lambda_\alpha \phi_{\alpha y_i}, \\ m_i z_i'' &= U_{z_i} + \Sigma_\alpha \lambda_\alpha \phi_{\alpha z_i}, \end{aligned} \tag{92}$$

where α has the range from 1 to m. In this and the following sections of this chapter sums will be indicated as usual and no umbral indices will be used.

Hamilton's principle is simply the statement that the differential equations (92) are the differential equations of the minimizing arcs of the problem of finding in the class of $3n$-dimensional arcs

$$x_i = x_i(t), \quad y_i = y_i(t), \quad z_i = z_i(t) \qquad (t_1 \leq t \leq t_2;\ i = 1, \cdots, n)$$

joining two given points and satisfying the equations $\phi_\alpha = 0$, one which minimizes the integral

$$I = \int_{t_1}^{t_2} (T + U)\, dt$$

where U is the force function and T the so-called kinetic energy

$$T = \tfrac{1}{2} \Sigma_i m_i v_i^2 = \tfrac{1}{2} \Sigma_i m_i (x_i'^2 + y_i'^2 + z_i'^2).$$

It is very easy to show that the equations (92) are the Euler-Lagrange equations for this problem. We have only to set up these equations for the function

$$F = T + U + \Sigma_\alpha \lambda_\alpha \phi_\alpha.$$

An important application of Hamilton's principle is that of determining the equations of motion in terms of the so-called generalized coördinates of Lagrange. The number of coördinates x_i, y_i, z_i is $3n$ and the number of equations $\phi_\alpha = 0$ is m. It is in general possible in an infinity of ways to express these coördinates as functions of t and $3n - m$ arbitrary parameters $q_1, \cdots, q_{3n-m}$ satisfying identically the equations $\phi_\alpha = 0$ and giving all the solutions of these equations. The functions T and U then take the form

$$T = T(t, q, q'), \quad U = U(t, q),$$

and the problem is transformed into that of finding among the arcs $q_r = q_r(t)$ $(r = 1, \cdots, 3n - m)$ joining the two given points one which minimizes the integral I. No adjoined conditions $\phi_a = 0$ are now necessary. The differential equations of the minimizing arcs for the new problem are the equations

$$\frac{d}{dt}\frac{\partial T}{\partial q_r'} - \frac{\partial}{\partial q_r}(T + U) = 0 \qquad (r = 1, \cdots, 3n - m).$$

The important result is that the form of these equations is the same no matter what new coördinates $q_1, \cdots, q_{3n-m}$ with the properties described above are used.

20. *Two forms of the principle of least action.** Let us now consider the somewhat special case where the functions U and ϕ_a of the last section do not contain the time t explicitly. If the equations (92) are multiplied by x_i', y_i', z_i', respectively, added, and integrated we find the well-known relation

$$T = U + h$$

where h is a constant of integration. This is the principle of the conservation of energy which says that the sum of the kinetic energy T and the potential energy $-U$ of a system satisfying equations (92) is always a constant.

Jacobi's form of the principle of least action states that the totality of dynamical trajectories satisfying equations (92) and having a given energy constant h is identical with the totality of extremals for the problem of finding among the arcs

$$x_i = x_i(u), \quad y_i = y_i(u), \quad z_i = z_i(u) \qquad (i = 1, \cdots, n;\ u_1 \leqq u \leqq u_2)$$

joining two given points and satisfying the equations $\phi_a = 0$ one which minimizes the integral

$$I = \int_{u_1}^{u_2} [2(U + h)S]^{1/2} du,$$

where S is simply a notation for the sum

$$S = \Sigma_i m_i (x_{iu}^2 + y_{iu}^2 + z_{iu}^2).$$

The parameter u is not in this case the time, but if at the time t_0 the particles are at the places defined on their trajectories by the parameter value u_0, then it turns out that the time at the place defined by u is

* Bolza [3, pp. 556, 586].

$$t = t_0 + \int_{u_0}^{u} \{S/[2(U+h)]\}^{1/2} du, \tag{93}$$

as one would expect from the relation $S(du/dt)^2 = 2T = 2(U+h)$.

To prove these statements we note that the function F for the minimizing problem just described is

$$F = [2(U+h)S]^{1/2} + \Sigma_a \mu_a \phi_a.$$

A typical one of the Euler-Lagrange equations is

$$(d/du)\{[2(U+h)]/S\}^{1/2} m_i x_{iu} - U_{x_i}\{S/[2(U+h)]\}^{1/2} - \Sigma_a \mu_a \phi_{a x_i} = 0.$$

If we introduce the parameter t along a solution of this equation by means of the formula (93) then the equation itself takes the form

$$m_i x_i'' - U_{x_i} - \Sigma_a \lambda_a \phi_{a x_i} = 0$$

when $\lambda_a = \mu_a\ (du/dt)$, which is the same as the first equation (92).

Lagrange's form of the principle of least action is again a principle for describing those mechanical trajectories which satisfy equations (92) and have a given energy constant h. They are extremals for the problem of finding among the arcs

$$x_i = x_i(t), \quad y_i = y_i(t), \quad z_i = z_i(t) \qquad (i = 1, \cdots, n;\ t_1 \leq t \leq t_2)$$

passing through given initial values of the coördinates for a given initial time t_1, passing through given end-values of the coördinates for an unspecified time t_2, and satisfying the equations

$$T - U - h = 0, \quad \phi_a = 0 \tag{94}$$

one which minimizes the integral

$$I = \int_{t_1}^{t_2} T dt.$$

This is a problem with a variable second end-point since t_2 is not specified. The function F for it is

$$F = T + \lambda(T - U - h) + \Sigma_a \mu_a \phi_a$$

and a typical Lagrange equation is

$$(d/dt)(1+\lambda) m_i x_i' + \lambda U_{x_i} - \Sigma_a \mu_a \phi_{a x_i} = 0. \tag{95}$$

When this equation is multiplied by x_i' and added to the other similar ones, it is found with the help of equations (94) that

$$\lambda = (k/2T) - 1/2$$

where k is a constant.

If all the end-values except x_2 are fixed in the theorem of pages 692-3, then the matrix (48) is square and its vanishing requires that

$$F(x_2) - \Sigma_i y_{i2}' F_{y_i'}(x_2) = 0.$$

Interpreted for the function F above this gives $\lambda = -1/2$ at $t = t_2$, with the help of equations (94). It follows that in the formula deduced above for λ we must have $k = 0$ and hence that $\lambda = -1/2$ for all values of t. Equation (95) then takes the form of the first equation (92) when we set $\lambda_a = 2\mu_a$.

CHAPTER III.

FURTHER NECESSARY CONDITIONS FOR A MINIMUM.

In this third chapter three further necessary conditions on a minimizing arc for the Lagrange problem will be developed, analogous to those of Weierstrass, Legendre, and Jacobi for the simpler types of problems of the calculus of variations. The analogue of Legendre's condition was first deduced by Clebsch [20] and the analogue of Jacobi's condition by A. Mayer [24]. For the deduction of these necessary conditions and for a number of other purposes we shall find the auxiliary theorems of the next section convenient.

21. *Two important auxiliary theorems.* Consider a one parameter family of admissible arcs

$$(96) \qquad y_i = y_i(x, b), \quad x_3(b) \leq x \leq x_4(b), \qquad (i = 1, \cdots, n)$$

for which the functions $x_3(b)$, $x_4(b)$, $y_i(x, b)$, $y_i'(x, b)$ are continuous and have continuous derivatives with respect to b in the domain of values (x, b) defined by the inequalities $b' \leq b \leq b''$, $x_3(b) \leq x \leq x_4(b)$, and whose end values describe two arcs C and D. The values of I taken along the arcs (96) are given by the formula

$$I(b) = \int_{x_3}^{x_4} f[x, y(x, b), y'(x, b)]dx$$

which has the derivative

$$I'(b) = f x_b \Big|_3^4 + \int_{x_3}^{x_4} \{f_{y_i} y_{ib} + F_{y_i'} y'_{ib}\} dx.$$

The index here is umbral and we shall use umbral indices freely elsewhere in this chapter. Since the arcs (96) are all admissible this result may also be written in the form

$$(97) \qquad \lambda_0 I'(b) = Fx_b \Big|_3^4 + \int_{x_3}^{x_4} \{F_{y_i} y_{ib} + F_{y_i'} y'_{ib}\} dx,$$

where the multipliers λ_0, $\lambda_a(x)$ in the function

$$F = \lambda_0 f + \lambda_a \phi_a$$

are entirely arbitrary. If now a particular arc of the family (96) satisfies the equations

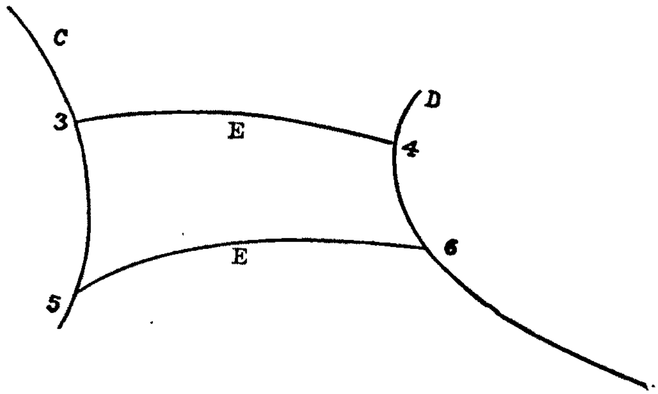
$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i$$

with a set of multipliers λ_0, $\lambda_a(x)$, then the introduction of these multipliers enables us to replace formula (97) by

$$\lambda_0 I'(b) = Fx_b + F_{y_i'} y_{ib} \Big|_3^4$$

where b is the particular value defining that arc. Since the equations of C and D are deduced from

$$x = x(b), \qquad y_i = y_i[x(b), b]$$



by replacing $x(b)$ by $x_3(b)$ and $x_4(b)$, respectively, it follows that along either of these arcs

$$dy_i = y'_i dx + y_{ib} db,$$

and therefore that

$$\lambda_0 dI = F dx + (dy_i - y_i' dx) F_{y_i'} \Big|_3^4.$$

Hence we have the following theorem:

AUXILIARY THEOREM. I. *Let*

$$(98) \qquad y_i = y_i(x, b), \quad x_1(b) \leq x \leq x_2(b), \qquad (i = 1, \cdots, n)$$

be a one-parameter family of admissible arcs without corners whose end-points describe two arcs C and D. If one of the arcs (98) *satisfies the equations*

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i \tag{99}$$

with a set of multipliers $\lambda_0, \lambda_\alpha(x)$ *then for the value of* b *defining it the values of* I *along the arcs* (98) *have a differential defined by the equation*

$$\lambda_0 dI = F dx + (dy_i - y_i' dx) F_{y_i'} \Big|_3^4. \tag{100}$$

In this formula the differentials dx, dy_i *at the point 3 are those of* C, *and at the point 4 those of* D.

If the particular arc along which the equation (99) holds is a normal arc then λ_0 can be taken equal to unity in formula (100). If each of the curves (98) has a set of multipliers $\lambda_0(b)$, $\lambda_\alpha(x, b)$ with which it satisfies equations (99), then the formula (100) holds along every arc of the family. We suppose that the functions $\lambda_0(b)$, $\lambda_\alpha(x, b)$ are continuous for $b' \leq b \leq b''$, $x_3(b) \leq x \leq x_4(b)$, and then we have

AUXILIARY THEOREM II. *Suppose that the arcs of the family* (98) *are all extremal arcs with multipliers of the form* $\lambda_0 = 1$, $\lambda_\alpha(x, b)$. *Then the values of* I *on two arcs* E_{34} *and* E_{56} *of the family satisfy the equation*

$$I(E_{56}) - I(E_{34}) = I^*(D_{46}) - I^*(C_{35})$$

with the values of the integral

$$I^* = \int \{F dx + (dy_i - y_i' dx) F_{y'_i}\}$$

along the corresponding segments C_{35} *and* D_{46} *shown in the last figure.*

This is readily found by integrating both sides of formula (100) with respect to b from the value b' defining simultaneously the points 3 and 4 to the value b'' defining similarly 5 and 6. The integrand of the integral I^* is readily seen to be a continuous function of b on the arcs C_{35} and D_{46} corresponding to the interval $b'b''$, on account of the properties of the functions $x(b)$, $y_i(x, b)$ defining the family (98).

22. *Necessary conditions analogous to those of Weierstrass and Legendre.* Suppose that the equations

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leq x \leq x_2)$$

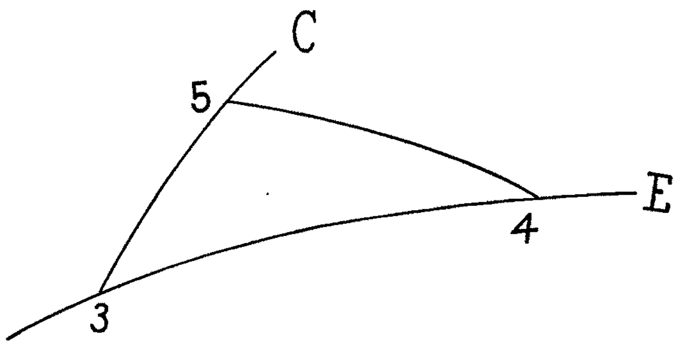
are those of a minimizing arc E_{12} for our problem.

We shall designate a set of values (x, y, y') as admissible if it lies in the neighborhood $\mathfrak{R}$ of page 1, satisfies the equations $\phi_\alpha = 0$, and gives the matrix

$\| \phi_{\alpha y_i'} \|$ the rank m. Let 3 be an arbitrary point on the arc E_{12} and let (x_3, y_{i3}, Y'_{i3}) be an admissible set. There is always an admissible arc

$$(C) \qquad y_i = Y_i(x) \qquad (x_3 \leq x \leq x_3 + h)$$

through this set since the equations $\phi_\alpha = 0$ determine uniquely m of the functions $Y_i(x)$ passing with their derivatives through the values prescribed by this initial set when the $n - m$ other functions $Y_i(x)$ have been chosen with initial values of themselves and their derivatives through their corresponding initial values of the set.



Suppose now that the arc E_{12} is normal on every sub-interval, and let 4 be so near to 3 on E_{12} that the arc E_{34} contains no corner. There is a $2n$-parameter family of admissible arcs $y_i = y_i(x, b_1, \cdots, b_{2n})$ containing E_{12} for $(b_1, \cdots, b_{2n}) = (0, \cdots, 0)$ and having $2n$ sets of variations $\eta_i(x)$ for which the determinant (38) with x_1, x_2 replaced by x_3, x_4 is different from zero. The $2n$ equations

$$y_i(x_5, b_1, \cdots, b_{2n}) = Y_i(x_5), \qquad y_i(x_4, b_1, \cdots, b_{2n}) = y_{i4}$$

have the initial solution $(x_5, b_1, \cdots, b_{2n}) = (x_3, 0, \cdots, 0)$ at which their functional determinant for $b_1, \cdots, b_{2n}$ is the determinant (38) with $x_1 = x_3$, $x_2 = x_4$ and different from zero. Hence they determine $2n$ functions $b_\mu = B_\mu(x_5)$ which vanish for $x_5 = x_3$. The family

$$y_i = y_i[x, B_1(x_5), \cdots, B_{2n}(x_5)] = y_i(x, x_5)$$

is now a one-parameter family of arcs joining the curve C of the figure to the point 4. The sum

$$\begin{aligned} \Phi(x_5) &= I(C_{35}) + I(E_{54}) \\ &= \int_{x_3}^{x_5} f(x, Y, Y')\,dx + \int_{x_5}^{x_4} f[x, y(x, x_5), y'(x, x_5)]\,dx \end{aligned}$$

must have its derivative ≥ 0 at x_3 if $I(E_{12})$ is to be a minimum. But with the help of formula (100) this derivative is seen to be

$$\Phi'(x_3) = E(x, y, y', Y', \lambda)\,|^3$$

if we define the E-function by the formula

$$E = F(x, y, Y', \lambda) - F(x, y, y', \lambda) - (Y_i' - y_i')F_{y_i'}(x, y, y', \lambda). \tag{101}$$

The multipliers in F are those associated uniquely with the normal minimizing arc E_{12}. Evidently one may always replace f by F for admissible sets (x, y, y'). We have then the following necessary condition:

ANALOGUE OF WEIERSTRASS NECESSARY CONDITION. *At each element (x, y, y', λ) of a minimizing arc which is normal on every sub-interval the inequality*

$$E(x, y, y', Y', \lambda) \geqq 0$$

must be satisfied for every admissible set $(x, y, Y') \neq (x, y, y')$.

The proof just given does not apply to the values x, y, y', λ at the right-hand end of an arc abutting on a corner, but it can be modified easily to be applicable by taking the point 4 at the left of 3, or one can infer the desired result by continuity considerations.

Consider now a set of values π_i satisfying the equations

$$\phi_{\alpha y_i'}\pi_i = 0 \tag{102}$$

at an element (x, y, y') of E_{12}. By means of the equations

$$\phi_{r y_i'}\pi_i = \kappa_r \tag{103}$$

these define $n - m$ further quantities κ_r. The equations

$$\phi_\alpha(x, y, p) = 0, \qquad \phi_r(x, y, p) = z_r + \epsilon\kappa_r$$

now have the initial solution $(\epsilon, p_1, \cdots, p_n) = (0, y_1', \cdots, y_n')$ and determine uniquely a set of solutions $p_i(\epsilon)$ with initial values $p_i(0) = y_i'$. The derivatives $p_i'(0)$ of these functions satisfy equations (102) and (103) when inserted in place of the numbers π_i and hence must coincide with them. The sets $(x, y, p(\epsilon))$ are now all admissible for sufficiently small values of ϵ, and according to the last theorem must satisfy the condition

$$E(x, y, y', p(\epsilon), \lambda) \geqq 0.$$

But we readily verify that this expression vanishes with its first derivative for ϵ at the value $\epsilon = 0$. Its second derivative

$$F_{y_i' y_k'}\pi_i\pi_k$$

at $\epsilon = 0$ must therefore be $\geqq 0$, from which we infer the

NECESSARY CONDITION OF CLEBSCH. *At every element* (x, y, y', λ) *of a minimizing arc which is normal on every sub-interval the inequality*

$$F_{y_i' y_k'}(x, y, y', \lambda)\pi_i\pi_k \geqq 0$$

must be satisfied by every set $(\pi_i, \cdot \cdot \cdot, \pi_n) \neq (0, \cdot \cdot \cdot, 0)$ *which is a solution of the m equations*

$$\phi_{\alpha y_i'}(x, y, y')\pi_i = 0.$$

23. *The envelope theorem.* According to the theorem of page 687, every extremal arc E_{12} along which the determinant R is different from zero is a member of a $2n$-parameter family of extremals of the form

$$y_i = y_i(x, a, b), \qquad \lambda_\alpha = \lambda_\alpha(x, a, b)$$

for special values a_{i0}, b_{i0} of the parameters. The family can be so chosen that the determinant (36) is different from zero at x_1, and we shall see in Section 27, page 727, that this determinant is in fact different from zero everywhere on E_{12}. If the constants a_i, b_i are replaced by functions $a_i(t)$, $b_i(t)$ with the initial values $a_i(0) = a_{i0}$, $b_i(0) = b_{i0}$ a one-parameter family of extremals is defined containing the arc E_{12} for the special parameter value $t = 0$. The arcs of this family will pass through the point 1 for $x = x_1$, and will touch an enveloping curve D at the points defined by a suitably chosen function $x(t)$, if the equations

$$\begin{aligned} &x' = k, \quad y_{ix}x' + y_{ia_k}a_k' + y_{ib_k}b_k' = ky_{ix}, \\ &y_{i1} = y_i(x_1, a, b) \end{aligned}$$

hold identically in t when x, a_i, b_i are replaced by the functions of t described above and the primes denote derivatives with respect to t. The first row of equations imposes the condition that the direction of the tangent to the curve D shall coincide with the direction $1 : y_1' : \cdot \cdot \cdot : y_n'$ of the tangent to the extremal. In order that these equations may be true it is evidently necessary and sufficient that the equations

$$\begin{aligned} &y_{ia_k}[x(t), a(t), b(t)]\, a_k' + y_{ib_k}[x(t), a(t), b(t)]\, b_k' = 0, \\ &y_{ia_k}[x_1, a(t), b(t)]\, a_k' + y_{ib_k}[x_1, a(t), b(t)]\, b_k' = 0, \end{aligned}$$

hold identically in t. If the derivatives a_k', b_k' are not zero it follows that the determinant

$$(104) \qquad \Delta(x, x_1, a, b) = \begin{vmatrix} y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{vmatrix}$$

vanishes identically in t when $x(t)$, $a_i(t)$, $b_i(t)$ are substituted.

DEFINITION OF A CONJUGATE POINT. A value $x_3 \neq x_1$ is said to define a *point 3 conjugate to 1 on the extremal arc E_{12}* if it is a root of a determinant $\Delta(x, x_1, a_0, b_0)$ belonging to a $2n$-parameter family of extremals $y_i = y_i(x, a, b)$, $\lambda_a = \lambda_a(x, a, b)$ for which the determinant

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

is different from zero on E_{12} as described on page 727.

Suppose now that 3 is such a conjugate point, and furthermore one at which the derivative Δ_x does not vanish. It is evident that if $\Delta_x \neq 0$ one at least of the minors of order $2n-1$ of Δ does not vanish at 3, and that the same property is therefore possessed by one at least of the determinants of order $2n$ of the matrix

$$\begin{Vmatrix} \Delta_x & \Delta_{a_k} & \Delta_{b_k} \\ 0 & y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ 0 & y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{Vmatrix}$$

since one at least of these determinants is the product of Δ_x by a non-vanishing minor of Δ. Then the first of the differential equations

$$\begin{aligned} \Delta_x(x, x_1, a, b)\,dx + \Delta_{a_k}(x, x_1, a, b)\,da_k + \Delta_{b_k}(x, x_1, a, b)\,db_k &= 0, \\ y_{ia_k}(x, a, b)\,da_k + y_{ib_k}(x, a, b)\,db_k &= 0, \\ y_{ia_k}(x_1, a, b)\,da_k + y_{ib_k}(x_1, a, b)\,db_k &= 0, \end{aligned} \tag{105}$$

with $2n-1$ of the others determine functions $x(t)$, $a_k(t)$, $b_k(t)$ with the initial values $x(0) = x_3$, $a_k(0) = a_{k0}$, $b_k(0) = b_{k0}$, and with derivatives x', a_k', b_k' not all zero at $t = 0$. Since $\Delta_x \neq 0$ at 3 it follows further that a_k', b_k' can not all vanish at $t = 0$. Since Δ vanishes at these initial values and has its derivative with respect to t identically zero, it must be itself identically zero in t. One sees readily then that the one remaining equation (105) is a consequence of the others when $x(t)$, $a_k(t)$, $b_k(t)$ are substituted. The following theorem is established:

Let E_{12} be an extremal arc along which the determinant R is different from zero, and let 3 be a point conjugate to 1 on E_{12} at which the derivative Δ_x of the determinant (104) *is different from zero. Then there exists through the point 1 a one-parameter family of extremals*

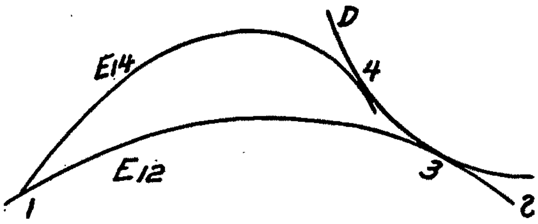
$$y_i = y_i(x, t), \qquad \lambda_a = \lambda_a(x, t) \tag{106}$$

containing E_{12} for the parameter value $t = 0$ and having an envelope D which touches E_{12} at the point 3. The functions y_i, y_{ix}, λ_a and the function $x(t)$

defining D have continuous derivatives in a neighborhood of the values x, t belonging to the arc E_{12}.

The last statement of the theorem is a consequence of the hypothesis (b) of page 676. For as a result of this hypothesis the functions y_i, y_{ix}, λ_a of the theorem on page 687 have continuous derivatives of the second order at least, and the solutions $x(t)$, $a_k(t)$, $b_k(t)$ of the equations (105) must therefore have continuous derivatives of at least the first order.

THE ENVELOPE THEOREM. *If the envelope D of the one-parameter family of extremals* (106) *has a branch projecting backward from 3 toward the*



point 1, as shown in the figure, then for every position of the point 4 on D preceding and near to 3 the arc $E_{14}+D_{43}+E_{32}$ is an admissible arc satisfying the equations $\phi_a=0$. Furthermore for every such arc

$$I(E_{14}+D_{43}+E_{32})=I(E_{12}).$$

Expressed in integral form the value of $I(E_{14}+D_{43})$ is

$$I(E_{14}+D_{43})=\int_{x_1}^{x(t)} f\,[x,y(x,t)y'(x,t)]\,dx+\int_t^{t_0} f\,x'\,dt$$

where the arguments in f in the last integral are $x(t), y[x(t),t], y'[x(t),t]$. The differential of the first integral with respect to t is given by formula (100) of page 716, and that of the second integral is readily found. It follows that

$$dI(E_{14}+D_{43})=-E(x,y,y',Y',\lambda)\,dx\,|^4$$

where Y' is the slope of D. But this vanishes identically in t since $Y'=y'$ at every point of D, and the final conclusion of the theorem is established. Evidently the envelope D satisfies the equations $\phi_a=0$ at each point 4 since it is tangent at that point to the extremal arc E_{14}.

24. *The analogue of Jacobi's condition.* The analogue of Jacobi's condition was discovered for the Lagrange problem by A. Mayer. Its statement is as follows:

4

THE NECESSARY CONDITION OF MAYER. *Let E_{12} be an extremal arc for the Lagrange problem which is normal on every sub-interval of x_1x_2 and has the determinant*

$$R = \begin{vmatrix} F_{y_i' y_k'} & \phi_{a y_i'} \\ \phi_{a y_k'} & 0 \end{vmatrix} \cdot$$

different from zero at every point of it. If E_{12} is a minimizing arc for the problem then between 1 and 2 on E_{12} there can be no point 3 conjugate to 1.

The proof of the statement for the case when the envelope has a branch as described in the envelope theorem is not difficult if one accepts the assertion that every extremal arc of a family $y_i(x, a, b)$ whose end-values x_1, x_2 and parameters a, b are sufficiently near to those of a normal extremal arc of the family is also normal. The proof of this assertion depends upon the fact that when the functions $y_i(x, a, b)$ are substituted in the equations of variation, the solutions $\eta_i(x, a, b)$ of those equations are continuous in the parameters a, b as well as x. Hence if there are $2n$ sets of variations η_{is} $(s = 1, \cdots, 2n)$ making the determinant (38) different from zero for the values x_{10}, x_{20}, a_0, b_0 defining the normal extremal, then this determinant will remain different from zero for neighboring values x_1, x_2, a, b.

If the arcs $E_{14} + D_{43} + E_{32}$ of the envelope theorem were all minimizing arcs they would necessarily have continuous multipliers since they have no corners. According to the assertion discussed in the last paragraph those sufficiently near to E_{12} would be normal on the intervals x_1x_4 and x_4x_2 since by hypothesis E_{12} is normal on every sub-interval and hence E_{13} and E_{32} are both normal. It follows readily that the composite arc $E_{14} + D_{43} + E_{32}$ would have the multipliers of the extremal E_{14} along E_{14}, the multipliers of the extremal tangent to D_{43} at each point of that arc, and the multipliers of the extremal E_{12} along E_{32}. Hence on the composite arcs near E_{12} the value of R would be everywhere different from zero as on E_{12}, and by the differentiability condition of page 684, each such arc would necessarily be an extremal. The extremal E_{12} is, however, the only one having its values y_i, v_i at $x = x_2$, or what is the same thing, its values y_i, y_i', λ_a at $x = x_2$. Hence the arcs $E_{14} + D_{43} + E_{32}$ can not all be minimizing arcs since otherwise all of them and the envelope D would necessarily fall upon E_{12} and their multipliers would coincide with those of E_{12}. But this is impossible because the derivatives $a_k'(t)$, $b_k'(t)$ of the family as determined on page 720 do not all vanish.

If an arc $E_{14} + D_{43} + E_{32}$ is not a minimizing arc it is always possible to find a neighboring admissible arc which joins the points 1 and 2 and gives the integral I a smaller value than $I(E_{14} + D_{43} + E_{32})$, that is, a smaller value than $I(E_{12})$, and hence $I(E_{12})$ can not be a minimum.

The preceding proof of the necessary condition of Mayer is a very satisfactory one geometrically because it emphasizes the geometrical interpretation of the conjugate point and the envelope theorem. But it rests upon two restrictive assumptions, namely, the non-vanishing of the derivative Δ_x at the conjugate point 3, and the requirement that the envelope have a branch projecting from 3 toward 1. In the following sections a proof of an entirely different sort is given which is free from these disadvantages.

25. *The second variation for a normal extremal.* It has been proved on page 17 that if the functions $\eta_i(x)$ of a set of admissible variations for a normal extremal arc E_{12} satisfy the relations $\eta_i(x_1) = \eta_i(x_2) = 0$, then there is a one-parameter family of admissible arcs

$$y_i = y_i(x, b) \qquad (x_1 \leq x \leq x_2)$$

joining the points 1 and 2, containing E_{12} for the parameter value $b = 0$, and having the functions $\eta_i(x)$ as its variations along E_{12}. When the various members of the equations

$$I(b) = \int_{x_1}^{x_2} f[x, y(x, b), y'(x, b)]\, dx,$$
$$0 = \phi_a[x, y(x, b), y'(x, b)]$$

are differentiated for b it is found that

$$I'(b) = \int_{x_1}^{x_2} (f_{y_i} y_{ib} + f_{y_i'} y_{ib}')\, dx,$$
$$0 = \phi_{ay_i} y_{ib} + \phi_{ay_i'} y_{ib}',$$

and a second differentiation gives for $b = 0$

$$I''(0) = \int_{x_1}^{x_2} (f_{y_i} y_{ibb} + f_{y_i'} y_{ibb}' + f_{y_i y_k} \eta_i \eta_k + 2 f_{y_i y_k'} \eta_i \eta_k' + f_{y_i' y_k'} \eta_i' \eta_k')\, dx,$$
$$0 = \phi_{ay_i} y_{ibb} + \phi_{ay_i'} y_{ibb}' + \phi_{ay_i y_k} \eta_i \eta_k + 2\phi_{ay_i y_k'} \eta_i \eta_k' + \phi_{ay_i' y_k'} \eta_i' \eta_k'.$$

When the last equations are multiplied by the factors λ_a, integrated from x_1 to x_2, and added to $I''(0)$ this derivative is found to have the value

$$I''(0) = \int_{x_1}^{x_2} (F_{y_i} y_{ibb} + F_{y_i'} y_{ibb}' + 2\omega)\, dx \tag{107}$$

where

$$2\omega(x, \eta, \eta') = F_{y_i y_k} \eta_i \eta_k + 2F_{y_i y_k'} \eta_i \eta_k' + F_{y_i' y_k'} \eta_i' \eta_k'. \tag{108}$$

On account of the equations

$$(d/dx) F_{y_i'} = F_{y_i}$$

the first two terms in the integral (107) have the anti-derivative $F_{y_i'}y_{ibb}$ and this vanishes at x_1 and x_2 as one readily sees by differentiating the equations

$$y_{i1} = y_i(x_1, b), \qquad y_{i2} = y_i(x_2, b)$$

twice with respect to b. Hence the following conclusions are justified:

Along a normal extremal arc E_{12} the second variation of the integral I is always expressible in the form

$$I''(0) = \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\,dx$$

where 2ω is the quadratic form defined by equation (108). *If $I(E_{12})$ is a minimum for the Lagrange problem then this second variation must be $\geqq 0$ for every set of admissible variations $\eta_i(x)$ whose functions satisfy the relations*

$$\eta_i(x_1) = \eta_i(x_2) = 0. \tag{109}$$

Since admissible variations satisfy the differential equations of variations

$$\Phi_a(x, \eta, \eta') = \phi_{ay_i}\eta_i + \phi_{ay_i'}\eta_i' = 0 \tag{110}$$

it is clear that these properties of the second variation suggest a minimum problem in $x\eta$-space of the same type as the original Lagrange problem in xy-space. There is an integral $I''(0)$ which must be $\geqq 0$ in the class of arcs $\eta_i = \eta_i(x)$ in $x\eta$-space satisfying the differential equations (110) and passing through the two fixed points $(x, \eta_1, \cdots, \eta_n) = (x_1, 0, \cdots, 0)$ and $(x, \eta_1, \cdots, \eta_n) = (x_2, 0, \cdots, 0)$, as indicated by equations (109). Evidently the minimum of $I''(0)$ in this class of arcs must be $\geqq 0$ if E_{12} is to be a solution of the original Lagrange problem.

The differential equations of the extremal arcs for the problem in $x\eta$-space are the equations

$$(d/dx)\Omega_{\eta_i'} = \Omega_{\eta_i}, \qquad \Phi_a(x, \eta, \eta') = 0 \tag{111}$$

where Ω is a function of the form

$$\Omega(x, \eta, \eta', \mu) = \mu_0\omega + \mu_a\Phi_a. \tag{112}$$

These are called by von-Escherich [31, Vol. 107, p. 1236] the *accessory system* of linear differential equations. They are the analogues of the Jacobi differential equation for the simplest problem in the plane. If the arc E_{12} is a normal extremal arc for the original Lagrange problem, then every extremal arc for the new problem in $x\eta$-space has this property, since the equations of variation of the linear equations $\Phi_a = 0$ for the $x\eta$-problem are these equations

themselves. Hence it is proper when E_{12} is normal to set $\mu_0 = 1$, the multipliers $\mu_0 = 1$, $\mu_a(x)$ for an extremal arc of the $x\eta$-problem being then unique.

The quadratic form $\Omega(x, \eta, \eta', \mu)$ has the properties

$$2\Omega = \eta_i \Omega_{\eta_i} + \eta_i' \Omega_{\eta_i'} + \mu_a \phi \mu_a, \tag{113}$$

$$u_i \Omega_{v_i} + u_i' \Omega_{v_i'} + \rho_a \Omega_{\sigma_a} = v_i \Omega_{u_i} + v_i' \Omega_{u_i'} + \sigma_a \Omega_{\rho_a}, \tag{114}$$

where the derivatives of Ω are understood to have the arguments (η, η', μ), (u, u', ρ), or (v, v', σ) as indicated by their subscripts. These are well-known formulas for quadratic forms which are readily provable and which will be useful in the following paragraphs.

A final remark concerning the accessory differential equations (111) is also important. These equations are linear and homogeneous in the variables $\eta_i, \eta_i', \eta_i'', \mu_a, \mu_a'$, and the determinant of coefficients of the variables η_i'', μ_a' is the determinant R which will be assumed different from zero along E_{12}. The arguments of Section 6 therefore tell us at once that the accessory equations have one and but one solution η_i, μ_a taking prescribed values of $\eta_i, \Omega_{\eta_i'}$ at a given value of x, or, what is the same thing, prescribed values of η_i, η_i', μ_a satisfying the equations of variation. In particular the only solution taking the values $\eta_i = \Omega_{\eta_i'} = 0$, or $\eta_i = \eta_i' = \mu_a = 0$, at a given x is the set of functions $\eta_i(x) \equiv \mu_a(x) \equiv 0$ which one readily sees to be a solution since the accessory equations are linear and homogeneous in $\eta_i, \eta_i', \eta_i'', \mu_a, \mu_a'$.

26. *A second proof of the analogue of Jacobi's condition.* Consider now a minimizing arc E_{12} for the original Lagrange problem, which has no corners and along which the determinant R of page 684 is everywhere different from zero. According to the differentiability condition on that same page the arc E_{12} must then be an extremal as defined in section 6. For the developments of the present section the additional assumption will be made that the extremal E_{12} is normal on every sub-interval of $x_1 x_2$.

DEFINITION OF CONJUGATE POINT. A value x_3 is said to define a *point 3 conjugate to 1 on the arc* E_{12} if there exists an extremal $\eta_i = u_i(x)$, $\mu_a = \rho_a(x)$ for the $x\eta$-problem whose functions $u_i(x)$ satisfy the relations $u_i(x_1) = u_i(x_3) = 0$ but are not identically zero on $x_1 x_3$. We shall presently see that the definition of a conjugate point on page 720 is equivalent to the one here given.

With this definition agreed upon the necessary condition of Mayer as stated on page 722 can be proved by showing that if there exists a point 3 conjugate to 1 between 1 and 2 on E_{12} then there exists also an admissible

set of variations $\eta_i(x)$ making $I''(0) < 0$. As a first step consider the functions $\eta_i(x)$, $\mu_a(x)$ defined by the equations

$$(115) \qquad \begin{aligned} \eta_i(x) &\equiv u_i(x), & \mu_a(x) &\equiv \rho_a(x) & &\text{on } x_1 \leq x \leq x_3, \\ \eta_i(x) &\equiv 0, & \mu_a(x) &\equiv 0 & &\text{on } x_3 \leq x \leq x_2, \end{aligned}$$

where the functions $u_i(x)$, $\rho_a(x)$ are those indicated in the definition just given for the conjugate point. With the help of the equations (112), (111), (113) it follows readily that for these functions $\eta_i(x)$

$$\begin{aligned} I''(0) &= \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\,dx = \int_{x_1}^{x_3} 2\Omega(x, u, u', \rho)\,dx \\ &= \int_{x_1}^{x_3} (u_i\Omega_{u_i} + u_i'\Omega_{u_i'} + \rho_a\Omega_{\rho_a})\,dx \\ &= u_i\Omega_{u_i'}\Big|_1^3 = 0. \end{aligned}$$

The functions $\eta_i(x)$ in (115) can not minimize $I''(0)$, however, since, as will be shown in the next paragraph, they do not satisfy the corner conditions

$$(116) \qquad \Omega_{\eta_i'}\,[x, \eta, \eta'(x-0), \mu(x-0)] = \Omega_{\eta_i'}\,[x, \eta, \eta'(x+0), \mu(x+0)]$$

at the point x_3. Hence there must be other admissible variations $\eta_i(x)$ vanishing at x_1 and x_2 and giving $I''(0)$ a value less than zero, and $I(E_{12})$ can not be a minimum.

To show that the corner conditions are not satisfied one may calculate readily the values of the derivatives $\Omega_{\eta_i'}$ for the functions (115) at the left and right of x_3. It is found then that the corner conditions (116) would require that $\Omega_{u_i'} = 0$ at the point x_3 as well as $u_i = 0$, and according to a remark at the end of the preceding section the functions $u_k(x)$, $\rho_a(x)$ would then have to be identically zero, which is not the case. The proof of Mayer's condition is now complete.

27. *The determination of conjugate points.* For a one-parameter family of extremals

$$y_i = y_i(x, b), \qquad \lambda_a = \lambda_a(x, b)$$

the equations

$$(d/dx)F_{y_i'} = F_{y_i}, \qquad \phi_a = 0$$

are identities in x and b. When they are differentiated with respect to b we find

$$\begin{gathered} (d/dx)(F_{y_i'y_k}y_{kb} + F_{y_i'y_k'}y_{kb}' + F_{y_i'\lambda_a}\lambda_{ab}) = F_{y_iy_k}y_{kb} + F_{y_iy_k'}y_{kb}' + F_{y_i\lambda_a}\lambda_{ab}, \\ \phi_{ay_k}y_{kb} + \phi_{ay_k'}y_{kb}' = 0, \end{gathered}$$

and these are precisely the accessory equations with the arguments $\eta_i = y_{ib}$, $\mu_a = \lambda_{ab}$. A $2n$-parameter family of extremals defined by equations similar to equations (35) or (37) on pages 686-7 furnishes by this differentiation process $2n$ solutions

$$(117)\qquad \begin{array}{l} y_{1a_k}, \cdots, y_{na_k};\ \lambda_{1a_k}, \cdots, \lambda_{ma_k} \\ y_{1b_k}, \cdots, y_{nb_k};\ \lambda_{1b_k}, \cdots, \lambda_{mb_k} \qquad (k=1, \cdots, n) \end{array}$$

of the accessory equations. The formulas of most importance here are those for $2n$-parameter families for which the determinant (36) is different from zero at some point, say x_1. We shall see in the next paragraph that it is then different from zero for all values of x.

Since the determinant R is different from zero along E_{12} the equations

$$\zeta_i = \Omega_{\eta_i'}(x, \eta, \eta', \mu), \qquad \Phi_a(x, \eta, \eta') = 0,$$

analogous to equations (30) on page 685, can be solved for η_k', μ_ρ. The solution has the form

$$(118)\qquad \eta_k' = G_k(x, \eta, \zeta), \qquad \mu_\beta = H_\beta(x, \eta, \zeta),$$

and the accessory equations are equivalent to the equations

$$(119)\qquad (d\eta_k/dx) = G_k(x, \eta, \zeta),\ (d\zeta_k/dx) = \Omega_{\eta_k}(x, \eta, G(x, \eta, \zeta), H(x, \eta, \zeta)).$$

All of these equations are linear and homogeneous in the arguments η_i, η_i', μ_a, ζ_i where they occur. For equations of the type (119) it is well known* that $2n$ solutions (η_k, ζ_k) whose determinant is different from zero at a single value of x, will have that determinant different from zero for all values x, and that every other solution is linearly expressible with constant coefficients in terms of $2n$ solutions which have this property. Every solution of the accessory equations is therefore expressible linearly with constant coefficients in terms of the $2n$ corresponding sets (η_k, μ_β) defined by the second of equations (118).*

Since the determinant (36) of page 687 is different from zero at $x = x_1$ it follows that it is different from zero for all values of x. For the $2n$ solutions (117) of the accessory equations define $2n$ solutions (η_k, ζ_k) of equations (119) whose determinant is different from zero. Hence every solution $(\eta_i, \mu_a) = (u_i, \rho_a)$ of the accessory equations is expressible in the form

$$u_i = c_k y_{ia_k} + d_k y_{ib_k}, \quad \rho_a = c_k \lambda_{aa_k} + d_k \lambda_{ab_k}.$$

---

* See, for example, Goursat, *A Course in Mathematical Analysis*, translated by Hedrick and Dunkel, Vol. 2, Part 2, pp. 153-4.

The values x_3 determining conjugate points according to the definition on page 725 are those for which the equations

$$u_i(x_3) = c_k y_{ia_k}(x_3) + d_k y_{ib_k}(x_3) = 0,$$
$$u_i(x_1) = c_k y_{ia_k}(x_1) + d_k y_{ib_k}(x_1) = 0,$$

have solutions c_k, d_k not all zero. But these are precisely the values x_3 for which the determinant $\Delta(x, x_1, a, b)$ vanishes, as indicated in the definition on page 720. We shall see on page 740 that for every ξ on x_1x_2 the zeros of $\Delta(x, \xi, a, b)$ are isolated from ξ when an extension of E_{12} is normal on every sub-interval.

Consider now an n-parameter family of extremals

$$y_i = y_i(x, b_1, \cdot\cdot\cdot, b_n), \qquad \lambda_a = \lambda_a(x, b_1, \cdot\cdot\cdot, b_n)$$

all of which pass through the point 1, and such that the functions $v_i = F_{y_i'}$ for the family have their determinant $| v_{ib_k} |$ different from zero at x_1. All of the derivatives y_{ib_k} vanish at x_1 as one may see by differentiating the equations

$$y_{i1} = y_i(x_1, b_1, \cdot\cdot\cdot, b_n)$$

with respect to b_k. Every solution η_i, μ_a of the accessory equations for which the functions $\eta_i(x)$ all vanish at x_1 is expressible in the form

$$\eta_i = c_k y_{ib_k}, \qquad \mu_a = c_k \lambda_{ab_k},$$

where the coefficients c_k are constants. For such a solution is uniquely determined by its set of values $\eta_i = 0$, $\zeta_i = \Omega_{\eta_i'}$ at $x = x_1$. If the constants c_k are solutions of the equations

$$\zeta_i(x_1) = c_k v_{ib_k}(x_1),$$

which in fact determine them uniquely, then the two solutions η_i, μ_a and $c_k y_{ib_k}$, $c_k \lambda_{ab_k}$ of the accessory equations have the same values $\eta_i = 0$, ζ_i at $x = x_1$ and hence are identical for all values of x. It follows that the points 3 conjugate to 1 on E_{12} are determined by values x_3 for which the equations

$$c_k y_{ib_k}(x_3) = 0$$

have solutions c_k not all zero, that is, by values $x_3 \neq x_1$ which make the determinant $D(x, b) = | y_{ib_k} |$ vanish. These results may be summarized as follows:

Let E_{12} be an extremal arc which is contained in a $2n$-parameter family of extremals

$$y_i = y_i(x, a_1, \cdots, a_n, b_1, \cdots, b_n), \quad \lambda_\alpha = \lambda_\alpha(x, a_1, \cdots, a_n, b_1, \cdots, b_n)$$

for special values a_{i0}, b_{i0} *of the parameters. Suppose furthermore that the determinant*

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

of the family, where $v_i = F_{y_i'}(x, y, y', \lambda)$, *is different from zero at the point 1 on* E_{12}. *Then the points 3 conjugate to 1 on* E_{12} *are determined by the roots* $x_3 \neq x_1$ *of the function* $\Delta(x, x_1, a_0, b_0)$ *where*

$$\Delta(x, x_1, a, b) = \begin{vmatrix} y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{vmatrix}.$$

If E_{12} *is a member of an n-parameter family of extremals*

$$y_i = y_i(x, b_1, \cdots, b_n), \quad \lambda_\alpha = \lambda_\alpha(x, b_1, \cdots, b_n)$$

all of which pass through the point 1, and such that the determinant $|v_{ib_k}|$ *for the functions* $v_i = F_{y_i'}$ *belonging to the family is different from zero at the point 1 on* E_{12}, *then the points conjugate to 1 on* E_{12} *are determined by the roots* $x_3 \neq x_1$ *of the function* $D(x, b_0)$ *where*

$$D(x, b) = |y_{ib_k}|$$

and the b_{i0} *are the parameter values defining* E_{12}.

CHAPTER IV.

SUFFICIENT CONDITIONS FOR A MINIMUM.

The conditions developed in the preceding chapters are conditions which must be satisfied by every minimizing arc for the Lagrange problem, but they have not been shown to actually insure the minimizing property. In this chapter it is proposed to discuss sets of conditions which are sufficient for a minimum. The methods of proof used are in essence those which Weierstrass applied in similar cases and which have been extended to the Lagrange problem by A. Mayer, Bolza, and others, but they involve important simplifications and improvements.

28. *Mayer fields and the fundamental sufficiency theorem.* The notion of a field has been defined in a number of different ways. The definition given here is not the usual one and is somewhat sophisticated, but it emphasizes properties which are well known for fields of the simplest problem in the

plane, and leads promptly to the theorem which is fundamental for all of the sufficiency proofs. In order to phrase this definition as simply as possible let us agree to call a set of values (x, y, y') *admissible* if it lies interior to the region $\mathfrak{R}$ where the continuity properties of the functions f and ϕ_a have been assumed, and satisfies the equations $\phi_a = 0$, and gives the matrix $\| \phi_{a y_i'} \|$ the rank m.

DEFINITION OF A MAYER FIELD. A *Mayer field* is a region $\mathfrak{F}$ of xy-space containing only interior points and having associated with it a set of functions

$$p_i(x, y), \quad l_a(x, y)$$

with the following properties:

(a) they have continuous first partial derivatives in $\mathfrak{F}$;

(b) the sets $(x, y, p(x, y))$ defined by the points (x, y) in $\mathfrak{F}$ are all admissible;

(c) the integral

$$I^* = \int \{F(x, y, p, l)dx + (dy_i - p_i dx)F_{y_i'}(x, y, p, l)\}$$

formed with these functions is independent of the path in $\mathfrak{F}$.

The integral I^* can also be written in the form

$$I^* = \int \{A dx + B_i dy_i\}$$

where

$$\begin{aligned} A(x, y) &= F(x, y, p, l) - p_k F_{y_k'}(x, y, p, l), \\ B_i(x, y) &= F_{y_i'}(x, y, p, l). \end{aligned}$$

If such an integral is independent of the path every arc is a minimizing arc for it and the Euler-Lagrange differential equations applied to it give the well-known conditions

$$\partial A/\partial y_i = \partial B_i/\partial x, \quad \partial B_i/\partial y_k = \partial B_k/\partial y_i \tag{120}$$

as necessary conditions for its invariantive property. One may readily prove the identities

$$\begin{aligned} \partial A/\partial y_i - \partial B_i/\partial x = F_{y_i} &- (\partial/\partial x)F_{y_i'} - p_k(\partial/\partial y_k)F_{y_i'} \\ &+ p_k(\partial B_i/\partial y_k - \partial B_k/\partial y_i) + \phi_a\, \partial l_a/\partial y_i \end{aligned} \tag{121}$$

where the partial derivatives indicated by the symbols ∂ are taken with respect to the independent variables x, y_i which occur explicitly and also in the field functions $p_i(x, y)$, $l_a(x, y)$.

From these results it is easy to see that in the field $\mathfrak{F}$ every solution $y_i(x)$ of the equations

$$dy_i/dx = p_i(x, y) \tag{122}$$

is an extremal with the multipliers $\lambda_a = l_a(x, y(x))$. For in the first place such an arc necessarily satisfies the equations $\phi_a = 0$, since the values (x, y, p) are all admissible; and in the second place the equations (120) and (121) then show that along such an arc

$$F_{y_i} - (d/dx)F_{y_i'} = F_{y_i} - (\partial/\partial x)F_{y_i'} - p_k(\partial/\partial y_k)F_{y_k'} = 0.$$

The arcs satisfying equations (122) are called the *extremals of the field.* Through each point of $\mathfrak{F}$ there passes one and but one such extremal arc since the equations (122) are of the first order. Furthermore the value of I^* along an extremal arc of the field is equal to that of the original integral I, since the equations $dy_i - p_i dx = 0$ are all satisfied along the field extremals.

If E_{12} is an extremal arc of a field $\mathfrak{F}$ then for every admissible arc C_{12} in the field joining the same two points 1 and 2 the formula

$$I(C_{12}) - I(E_{12}) = \int_{x_1}^{x_2} E[x, y, p(x, y), y', l(x, y)]\,dx \tag{123}$$

holds, where

$$E = F(x, y, y', l) - F(x, y, p, l) - (y_i' - p_i)F_{y_i'}(x, y, p, l)$$

and the arguments $y(x)$, $y'(x)$ in the integrand are those belonging to C_{12}.

The formula (123) is the analogue of a well-known one of Weierstrass and the proof of it is very simple. For since I^* is independent of the path in $\mathfrak{F}$ and has the same values as I along an extremal of the field it follows that

$$I(E_{12}) = I^*(E_{12}) = I^*(C_{12}),$$

and hence that

$$I(C_{12}) - I(E_{12}) = I(C_{12}) - I^*C_{12}).$$

The last two terms give the integral in the second member of the formula (123) when the integrand f in $I(C_{12})$ is replaced by F. This is evidently permissible since C_{12} is by hypothesis an admissible arc and therefore satisfies the equations $\phi_a = 0$.

With these results in mind it is now possible to prove the following important theorem:

THE FUNDAMENTAL SUFFICIENCY THEOREM. *If E_{12} is an extremal arc of a field $\mathfrak{F}$ and if at each point of the field the condition*

$$E[x, y, p(x, y), y', l(x, y)] > 0$$

holds for every admissible set (x, y, y') *different from* (x, y, p), *then the inequality* $I(C_{12}) > I(E_{12})$ *is true for every admissible arc* C_{12} *in the field and joining the end-points of* E_{12} *but not identical with* E_{12}.

It is evident from formula (123) that the inequality $I(C_{12}) \geqq I(E_{12})$ is necessarily satisfied. The equality sign is appropriate only if the E-function vanishes at every point of C_{12}, that is, only if the equations $y_i' = p_i$ are satisfied at each point of C_{12}. But in that case the arc C_{12} would coincide with E_{12} since the equations $y_i' = p_i$ have only one solution through the point 1 and that is E_{12} itself.

29. *The construction of a field.* The extremal arcs of a field may be regarded as forming an n-parameter family since one of them passes through each point of the field. By analogy with the properties of fields for the simplest problem of the calculus of variations in the plane it might be expected that every n-parameter family of extremals which simply covers a region in xy-space would provide a set of slope functions and multipliers $p_i(x, y)$, $l_\alpha(x, y)$ which would make the integral I^* independent of the path in that region, and hence form a field over the region, but such is not the case. The n-parameter families which can form fields are special in character in somewhat the same way that a two-parameter family of straight lines in xyz-space is special if it is cut orthogonally by a surface. It is well known that not every such family of straight lines has an orthogonal surface.

Let the equations

$$(124) \qquad y_i = y_i(x, a_1, \cdots, a_n), \quad \lambda_\alpha = \lambda_\alpha(x, a_1, \cdots, a_n)$$

be an n-parameter family of extremals with the property that the functions y_i, y_{ix}, λ_α have continuous first partial derivatives for all values $(x, a_1, \cdots, a_n)$ satisfying conditions of the form

$$(125) \qquad \begin{gathered} \xi_1(a_1, \cdots, a_n) \leq x \leq \xi_2(a_1, \cdots, a_n). \\ (a_1, \cdots, a_n) \text{ in a region } A. \end{gathered}$$

Suppose further that there is an n-space

$$x = x_1(a_1, \cdots, a_n), \quad y_i = y_i(x_1(a_1, \cdots, a_n), a_1, \cdots, a_n)$$

cutting the extremals (124) for which the function $x_1(a_1, \cdots, a_n)$ has continuous first partial derivatives in A. The extremals (124) are said to simply cover a field $\mathfrak{F}$ of points (x, y) if to each point of the region there corresponds one and but one set of values x, $a_i(x, y)$ satisfying the first n equations (124)

and the conditions (125); and if the functions $a_i(x,y)$ so defined have continuous derivatives in $\mathfrak{F}$. The functions

$$p_i(x,y) = y_{ix}[x, a(x,y)], \quad l_a(x,y) = \lambda_a[x, a(x,y)]$$

are then a set of slope-functions and multipliers for the region $\mathfrak{F}$, and the following theorem can be proved:

Suppose that an n-parameter family of extremals

$$y_i = y_i(x, a_1, \cdots, a_n), \quad \lambda_a = \lambda_a(x, a_1, \cdots, a_n) \tag{126}$$

is intersected by an n-space

$$x = x_1(a_1, \cdots, a_n), \quad y_i = y_i(x_1(a_1, \cdots, a_n), a_1, \cdots, a_n) \tag{127}$$

and simply covers a region $\mathfrak{F}$ of xy-space containing only interior points, in the manner described in the preceding paragraphs. If the parameter values of the extremal through a point (x,y) are denoted by $a_i(x,y)$ then the region $\mathfrak{F}$ is a field with the slope-functions and multipliers

$$p_i(x,y) = y_{ix}[x, a(x,y)], \quad l_a(x,y) = \lambda_a[x, a(x,y)] \tag{128}$$

provided that the integral I^* is independent of the path in the n-space (127).

The proof may be made with the help of the Auxiliary Theorem II of page 716. For an arc D_{46} in $\mathfrak{F}$ with equations of the form

$$x = x(t), \quad y_i = y_i(t) \qquad (t' \leq t \leq t'')$$

defines a one-parameter family of extremals intersecting it, and a corresponding arc C_{35} in the n-space (127), by means of the functions $a_i(t) = a_i[x(t), y(t)]$. According to the auxiliary theorem cited it is then true that

$$I^*(D_{46}) = I^*(C_{35}) + I(E_{56}) - I(E_{34}).$$

The three terms on the right are completely determined when the end-points of D_{46} are given, since by hypothesis the value $I^*(C_{35})$ is the same for all arcs C_{35} with the same end-points in the n-space (127). Hence the integral I^* is independent of the path in the whole of the region $\mathfrak{F}$, as required by the definition of a field.

The preceding theorem suggests at once a number of methods of constructing fields by means of n-parameter families of extremals. One may take the n-parameter family through a fixed point O and regard the point O as a degenerate n-space (127). Certainly on this degenerate n-space the integral I^* is independent of the path. Every region in xy-space simply covered by

the extremals will then be a field with the slope-functions and multipliers (128).

If an n-space

$$(129) \qquad x = X(a_1, \cdots, a_n), \quad y_i = Y_i(a_1, \cdots, a_n)$$

and a function $W(a_1, \cdots, a_n)$ are chosen arbitrarily in advance the $n + m$ equations

$$(130) \qquad FX_{a_i} + (Y_{ka_i} - y_k' X_{a_i}) F_{y_k'} = W_{a_i}, \quad \phi_\alpha = 0,$$

where the arguments of F, ϕ_α are X, Y_i, y_i', λ_α, may under certain conditions be solved for the $n + m$ variables y_i', λ_α as functions of $a_1, \cdots, a_n$. At each point of the n-space an initial element x, y_i, y_i', λ_α of an extremal is thus determined, and the extremals which have these initial elements form an n-parameter family. The integrand of the integral I^* for this family has the value dW on every arc in the n-space (129), on account of the equations (130), since along such an arc the differentials dx, dy_k have the values

$$dx = X_{a_i} da_i, \quad dy_k = Y_{ka_i} da_i.$$

Hence the integral I^* will be independent of the path on the space (129) and every region of xy-space simply covered by the family of extremals will form a field. If the derivatives W_{a_i} all vanish then an n-space (129) which satisfies the equations (130) with the extremals of the family it is said to cut the family transversally.

A similar discussion can be made for initial spaces (129) of lower dimensions.

30. *Sufficient conditions for a strong relative minimum.* In the following paragraphs the necessary conditions deduced in the preceding chapters will be designated by the numerals I, II, III, IV. These are, respectively, the necessary condition of page 683, the analogue of Weierstrass' condition on page 718, the condition of Clebsch on page 719, and the condition of Mayer on page 722. The notations II′, III′ will be used to designate the conditions II and III when strengthened to exclude the equality sign which occurs in their statements. Similarly IV′ is the stronger condition of Mayer which excludes the conjuate point 3 from the end-point 2 of E_{12}, as well as from the interior of that arc. An arc E_{12} with multipliers $\lambda_0 = 1$, $\lambda_\alpha(x)$ will be said to satisfy the condition II$_b$′ if the inequality

$$E(x, y, y', Y', \lambda) > 0$$

holds for every set of elements (x, y, y', Y', λ) for which the set (x, y, y', λ)

is in a neighborhood of similar sets belonging to E_{12}, and $(x, y, Y') \neq (x, y, y')$ is admissible.

Every extremal arc E_{12} defined on an interval $x_1 x_2$, and on which the determinant R is different from zero, defines an extended extremal on an interval $x_1 - d \leqq x \leqq x_2 + d$ which contains E_{12} as part of it. We may call this longer extremal an extension of E_{12}.

With these agreements we can state the following theorem:

SUFFICIENT CONDITIONS FOR A STRONG RELATIVE MINIMUM. *If an admissible arc E_{12}, without corners and with an extension normal on every subinterval, satisfies the conditions I, II_b', III′, IV′, then there is a neighborhood $\mathfrak{F}$ of the points (x, y) on E_{12} such that the inequality $I(C_{12}) > I(E_{12})$ holds for every admissible arc C_{12} which is in $\mathfrak{F}$ and not identical with E_{12}.*

The minimum furnished by E_{12} is called a relative minimum because it is in a class of arcs restricted to lie in a neighborhood $\mathfrak{F}$ of E_{12}; and it is a strong relative minimum because the neighborhood $\mathfrak{F}$ lays no restriction on the slopes y_i' of comparison arcs which lie in it.

In order to prove the theorem we should note in the first place that the condition I and the normality of E_{12} imply a unique set of multipliers $\lambda_0 = 1$, $\lambda_\alpha(x)$ and constants c_i with which E_{12} satisfies the equations (24) of page 683.

The condition III' now implies that the determinant R of page 684 is different from zero at every element (x, y, y', λ) of E_{12}. For at an element where R vanished the linear equations

$$(131) \qquad F_{y_i' y_k'} \Pi_k + \phi_{\alpha y_i'} \mu_\alpha = 0, \quad \phi_{\alpha y_k'} \Pi_k = 0$$

would have solutions Π_k, μ_α not all zero, with the numbers Π_k also not all zero since the matrix $\| \phi_{\alpha y_i'} \|$ has rank m. But when the first equations (131) are multiplied by $\Pi_1, \cdots, \Pi_n$ and added it is found that

$$F_{y_i' y_k'} \Pi_i \Pi_k = 0,$$

as a result of the second set of equations (131), which would contradict the condition III′. .

Since the determinant R is different from zero along E_{12} it follows from the differentiability condition of page 684 that E_{12} must be an extremal. According to the developments of Section 6, page 687, there exists a $2n$-parameter family of extremals

$$y_i = y_i(x, a, b), \qquad \lambda_\alpha = \lambda_\alpha(x, a, b)$$

containing E_{12} for special parameter values a_{i0}, b_{i0}. The functions y_i, y_{ix}, λ_α have continuous partial derivatives of the first three orders near the values

(x, a_i, b_i) belonging to E_{12}, and the determinant (36) of page 687 is different from zero at the point 1 on E_{12}.

It will be shown in Section 32 that for an arc E_{12} with an extension normal on every sub-interval there is always an interval $x_1 - h \leq x \leq x_1 + h$ containing no pair of conjugate points, or in other words, containing no two values x, x_0 which satisfy the equation $\Delta(x, x_0, a_0, b_0) = 0$, where Δ is the determinant (104) of page 719. Hence if $x_0 < x_1$ be chosen sufficiently near to x_1 the function $\Delta(x, x_0, a_0, b_0)$ will be different from zero on the interval $x_1 \leq x \leq x_1 + h$, and different from zero also in the interval $x_1 + h \leq x \leq x_2$ on account of the continuity of Δ and condition IV′. The equations

$$(132) \qquad y_i = y_i(x, a, b), \qquad y_{i0} = y_i(x_0, a, b)$$

have now as initial solutions the totality of values (x, y, a, b) belonging to E_{12}, and their functional determinant $\Delta(x, x_0, a, b)$ with respect to the parameters a_i, b_i is different from zero at these initial solutions on account of the choice of x_0 which has just been made. Well-known implicit function theorems then justify the statement that there is a neighborhood $\mathfrak{F}$ of the points (x, y) on E_{12} in which the equations (132) have solutions $a_i(x, y)$, $b_i(x, y)$ with continuous partial derivatives of the first three orders since the functions (132) have such derivatives. This neighborhood $\mathfrak{F}$ is a field with the slope functions and mutlipliers

$$p_i(x, y) = y_{ix}[x, a(x, y), b(x, y)], \quad \lambda_a(x, y) = \lambda_a[x, a(x, y), b(x, y)]$$

since the extremals which simply cover it all pass through the fixed point 0 corresponding, on E_{12} extended, to the value x_0. If the field $\mathfrak{F}$ is taken sufficiently small the values $x, y, p_i(x, y), \lambda_a(x, y)$ belonging to it will remain in so small a neighborhood of the sets (x, y, y', λ) belonging to E_{12} that according to the condition II_b' the inequality

$$(133) \qquad E[x, y, p(x, y), y, \lambda(x, y)] > 0$$

will hold for every admissible element $(x, y, y') \neq (x, y, p)$ in $\mathfrak{F}$. The fundamental sufficiency theorem then justifies the theorem which was to be proved.

31. *Sufficient conditions for a weak relative minimum.* The conditions I, III′, IV′ were the only ones used in the last section up to the very last paragraph. If they only are assumed it is not possible to establish the condition (133). The E-function for admissible elements (x, y, y') in the field $\mathfrak{F}$ is expressible, however, with the help of Taylor's formula with integral remainder term, in the form

$$(134) \qquad E = (y_i' - p_i)(y_k' - p_k) \int_0^1 (1-\theta) F_{y_i' y_k'}[x, y, p + \theta(y' - p), \lambda] \, d\theta$$

where $p_i = p_i(x, y)$, $\lambda_a = \lambda_a(x, y)$ are the slope-functions and multipliers of the field, and the differences $y_i' - p_i$ satisfy the equation

$$\phi_a(x, y, y') - \phi_a(x, y, p) = (y_i' - p_i) \int_0^1 \phi_{ay_i'}[x, y, p + \theta(y' - p)] d\theta = 0.$$

On account of the condition III′ the quadratic form

$$\Pi_i \Pi_k \int_0^1 (1 - \theta) F_{y_i' y_k'}[x, y, p + \theta(y' - p), \lambda] d\theta$$

is positive for all sets (x, y, y', Π) for which (x, y, y') is on the arc E_{12} where $y_i' = p_i(x, y)$, and for which the numbers Π_i satisfy the equations

$$\Pi_i \Pi_i = 1, \quad \Pi_i \int_0^1 \phi_{ay_i'}[x, y, p + \theta(y' - p)] d\theta = 0.$$

Hence it stays positive for sets of values (x, y, y', Π) for which the numbers Π_i satisfy these equations and the set (x, y, y') lies in a sufficiently small neighborhood N of similar sets on E_{12}. It follows readily that the E-function (134) of the field $\mathfrak{F}$ is positive at least for all sets $(x, y, y') \neq (x, y, p)$ in the neighborhood N, and the following theorem is therefore justified:

SUFFICIENT CONDITIONS FOR A WEAK RELATIVE MINIMUM. *If an admissible arc E_{12} without corners and with an extension normal on every sub-interval, satisfies the conditions I, III′, IV′ then there is a neighborhood N of the sets of values (x, y, y') on E_{12} such that the inequality $I(C_{12}) > I(E_{12})$ holds for every admissible arc C_{12} whose elements (x, y, y') are all in N but which is not identical with E_{12}.*

The minimum described in this theorem is called a weak relative minimum because the neighborhood N in which it exists requires the slopes y_i' of the comparison arcs C_{12}, as well as their points (x, y), to be near those on E_{12}.

32. *The justification of a preceding statement.* It was stated on page 736 that there is always an interval $x_1 - h \leq x \leq x_1 + h$ on which no two values x, x_0 can satisfy the equation $\Delta(x, x_0, a_0, b_0) = 0$. The proof of this statement is not simple, but it can be made with the help of properties of solutions of the accessory differential equations

$$(135) \qquad (d/dx)\Omega_{\eta_i'} - \Omega_{\eta_i} = 0, \quad \Omega_{\mu_a} = \Phi_a = 0$$

for the arc E_{12} described on page 724. It is understood that the arc E_{12} is an extremal with an extension normal on every sub-interval and satisfying the condition III′. As a consequence of these properties the determinant R is different from zero at every point of E_{12}.

The equation (114) on page 725

$$u_i\Omega_{v_i} + u_i'\Omega_{v_i'} + \rho_a\Omega_{\sigma_a} = v_i\Omega_{u_i} + v_i'\Omega_{u_i'} + \sigma_a\Omega_{\rho_a},$$

justifies readily the further relation

$$u_i[\Omega_{v_i} - (d/dx)\Omega_{v_i'}] + \rho_a\Omega_{\sigma_a} - v_i[\Omega_{u_i} - (d/dx)\Omega_{u_i'}] - \sigma_\sigma\Omega_{\rho_a} = (d/dx)(v_i\Omega_{u_i'} - u_i\Omega_{v_i'}).$$

Hence for every pair of solutions u_i, ρ_a and v_i, σ_a of the accessory equations the expression

$$\psi(u, \rho, v, \sigma) = u_i\Omega_{v_i'} - v_i\Omega_{u_i'}$$

is a constant. If this constant is zero the two solutions are said to be *conjugate solutions.*

There is one and but one set of solutions η_i, μ_a of the accessory equations (135) for which η_i, $\zeta_i = \Omega_{\eta_i'}$ take assigned values at the value x_1, as shown for the original xy-problem on pages 685 and 686. A matrix of n solutions u_{ik}, ρ_{ak} $(k = 1, \cdots, n)$ therefore exists for which at the value x_1 the matrix $\| u_{ik} \|$ is the identity matrix and the corresponding matrix of the functions $\zeta_i = \Omega_{\eta_i'}$ has all its elements zero. The solutions u_{ik}, ρ_{ak} $(k = 1, \cdots, n)$ are conjugate in pairs, as one readily verifies, since their functions ζ_i all vanish at x_1. The notations u_i, ρ_a and v_i, σ_a will be used for the linear expressions

$$u_i = a_k u_{ik}, \qquad \rho_a = a_k\rho_{ak},$$
$$v_i = a_k' u_{ik}, \qquad \sigma_a = a_k'\rho_{ak},$$

where the coefficients a_k are functions of x to be determined and the variables a_k' are derivatives of the coefficients a_k with respect to x. Primes attached to expressions involving u_i, ρ_a or v_i, σ_a will always indicate derivatives of those expressions with respect to x calculated as if the coefficients a_k, a_k' were independent of x. One readily verifies, then, the relations

$$(136) \qquad \begin{gathered} (\Omega_{u_i'})' = \Omega_{u_i}, \quad (\Omega_{v_i'})' = \Omega_{v_i}, \quad u_i\Omega_{v_i'} - v_i\Omega_{u_i'} = 0, \\ (d/dx)\Omega_{u_i'} = (\Omega_{u_i'})' + \Omega_{v_i'} = \Omega_{u_i} + \Omega_{v_i'} \end{gathered}$$

in which it is understood that the differentiation indicated by d/dx takes account of the fact that the coefficients a_i are functions of x.

Let the functions $\eta_i(x)$ be a set of admissible variations along the arc E_{12}, satisfying therefore the equations $\Phi_a = 0$. The equations

$$\eta_i = u_i = a_k u_{ik}, \quad \mu_a = \rho_a = a_k\rho_{ak}$$

determine uniquely the coefficients a_k and the multipliers μ_a as functions of x on an interval $x_1 - h \leq x \leq x_1 + h$ chosen so small that on it the determinant $| u_{ik} |$ is everywhere different from zero. The derivatives η_i' have the values

$$\eta_i' = a_k u_{ik}' + a_k' u_{ik} = u_i' + v_i. \tag{137}$$

With the help of Taylor's formula, equation (113) of page 725, the equations (136) and (137) above, and the relations $\Omega_{\rho_a} = \Phi_a = 0$, one verifies the further relations

$$\begin{aligned} 2\omega(x, \eta, \eta') &= 2\Omega(x, \eta, \eta', \rho) = 2\Omega(x, u, u' + v, \rho) \\ &= 2\Omega(x, u, u', \rho) + 2v_i\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= u_i\Omega_{u_i} + u_i'\Omega_{u_i'} + \rho_a\Omega_{\rho_a} + 2v_i\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= u_i[\Omega_{u_i} + \Omega_{v_i'}] + (u_i' + v_i)\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= (d/dx)(\eta_i\Omega_{u_i'}) + F_{y_i' y_k'}(\eta_i' - u_i')(\eta_k' - u_k'). \end{aligned}$$

For arbitrary multipliers $\mu_a(x)$ taken with the functions $\eta_i(x)$ it follows therefore that

$$2\Omega(x, \eta, \eta', \mu) = (d/dx)\eta_i\Omega_{u_i'} + F_{y_i' y_k'}(\eta_i' - u_i')(\eta_k' - u_k')$$

and hence with the help of equation (113) on page 725 that

$$\eta_i[\Omega_{\eta_i} - (d/dx)\Omega_{\eta_i'}] + (d/dx)\eta_i(\Omega_{\eta_i'} - \Omega_{u_i'}) = F_{y_i' y_k'}(\eta_i' - u_i')(\eta_k' - u_k').$$

The last equation justifies the following lemma:

LEMMA. *There is an interval $x_1 - h \leq x \leq x_1 + h$ on which there exists no solution $\eta_i(x)$, $\mu_a(x)$ of the accessory equations, except the solution $\eta_i \equiv \mu_a \equiv 0$, whose elements $\eta_i(x)$ all vanish at two points x' and x'' of the interval; or, in other words, there is an interval on which no pair of values x', x'' can define conjugate points on E_{12}.*

This is clear since the last equation shows that for a system of solutions $\eta_i(x)$, $\mu_a(x)$ of the accessory equations the sum $\eta_i(\Omega_{\eta_i'} - \Omega_{u_i'})$ has a non-negative derivative on $x_1 - h \leq x \leq x_1 + h$, on account of the property III′ of E_{12}. If the functions $\eta_i(x)$ all vanish at two points x' and x'' the differences $\eta_i' - u_i' = v_i$ are identically zero on $x'x''$, and this implies that the derivatives a_k' are all zero and the coefficients a_k constants. But since the $\eta_i(x)$ vanish at x' and $| u_{ik} |$ is different from zero these coefficients are then all zero, and the functions $\eta_i(x)$ vanish identically on $x'x''$. The multipliers $\mu_a(x)$ are also zero on $x'x''$. Otherwise they would form with $\lambda_0 = 0$ a set of multipliers for E_{12}, as one readily sees by examining the accessory equations, and this is impossible since the extension of E_{12} is normal on $x'x''$ if the interval $x_1 - h \leq x \leq x_1 + h$ is taken sufficiently small.

As an immediate consequence of this lemma we have the following corollary:

Corollary. *There is an interval $x_1 - h \leqq x \leqq x_1 + h$ on which the determinant*

$$\Delta(x, x_0, a, b) = \begin{vmatrix} y_{ia_k}(x) & y_{ib_k}(x) \\ y_{ia_k}(x_0) & y_{ib_k}(x_0) \end{vmatrix},$$

formed for a family of extremals $y_i = y_i(x, a, b)$, $\lambda_a(x, a, b)$ as described in the theorem of page 687, can not vanish for any pair of points $(x_0, x) = (x', x'')$.

The solutions η_i, μ_a of the accessory equations are all expressible in the form

$$(138) \qquad \eta_i = c_k y_{ia_k} + d_k y_{ib_k}, \qquad \mu_a = c_k \lambda_{aa_k} + d_k \lambda_{ab_k},$$

as was indicated on page 727. If $\Delta(x'', x', a, b) = 0$ for points x', x'' on the interval $x_1 - h \leqq x \leqq x_1 + h$ then there would be constants c_k, d_k not all zero such that the solution (138) has $\eta_i(x') = \eta_i(x'') = 0$, and by the lemma it would follow that $\eta_i \equiv \mu_a \equiv 0$. In that case the corresponding functions

$$\zeta_i = c_k v_{ia_k} + d_k v_{ib_k} = \Omega_{\eta_i'}$$

would also vanish identically, which is impossible since the determinant

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

of page 687 is by hypothesis different from zero.

CHAPTER V.

Historical Remarks.

A complete history of the problem of Lagrange would require an extensive presentation. The remarks in the following paragraphs are a sketch only of the development of the theory, in which an effort will be made to point out the memoirs which have been especially significant in the preparation of this paper. For more detailed references one should consult the articles on the calculus of variations in the *Encyclopädie der Mathematischen Wissenschaften* by Kneser [1, II A 8] * and Zermelo and Hahn [1, II A 8 a], the translations and extensions of them by Lecat in the *Encyclopédie des Sciences Mathématiques* [2], and the treatise by Bolza [3].

* The numbers in square brackets refer to the following bibliography.

Euler [2, p. 119; 7, p. 114] and Lagrange [8, I, p. 347] both studied special cases of the Lagrange problem which led up to the formulation of the more general problem and its multiplier rule by Lagrange [8, X, p. 420]. The proof of the multiplier rule which Lagrange gave was incomplete. The missing details were provided by A. Mayer [9], Hilbert [10], and Kneser [11, Sections 57-8]. Hahn [12] extended to the multiplier rule for the problem of Mayer, which includes that of Lagrange as a special case, the methods which Du Bois Reymond had applied to simpler problems of the calculus of variations. The argument in the text above is new but was suggested by papers by Hahn [13, p. 271] and Bliss [16].

The distinction between normal and abnormal minimizing arcs seems to have been first mentioned by A. Mayer [9, p. 79] but was emphasized by von Escherich [17] in connection with his theory of the second variation where it played an important role. Hahn [18, p. 152] adopts the definition of von Escherich. The definitions in Sections 7 and 8 above are modeled after that of Bolza [19, p. 440] and are applied to simplify the proof of the multiplier rule in Section 15 for the case when the functions ϕ_a contain no derivatives.

The necessary condition analogous to that of Legendre for simpler problems was first proved for the problem of Lagrange by Clebsch [20] as one of the consequences of his rather elaborate theory of the second variation. The necessary condition analogous to that of Weierstrass seems to have been first proved by Hahn [21] who deduced therefrom the necessary condition of Clebsch without appeal to the theory of the second variation. The method in the text above is that of Bolza [22], who supplied a step missing in the proof of Hahn, but the method is here further simplified by the use of the auxiliary formulas of Section 21 which are generalizations of formulas emphasized by Goursat [23, p. 566].

For the Lagrange problem the necessary condition for a minimum analogous to that of Jacobi for simpler problems is due to A. Mayer [24]. The envelope theorem and the associated geometric proof of the Mayer condition are the work of Kneser [25]. The method of the preceding pages for the development of Kneser's theory is modeled after Bolza [26], but with simplifications due again to the use of the auxiliary formulas of Section 21. The analytic proof of the Mayer condition by means of the theory of the minimum problem of the second variation was suggested by Bliss [27] and applied to the Lagrange problem by D. M. Smith [28]. By this method the advantages of the analytic proof are preserved without the necessity of using any complicated theory of the transformation of the second variation.

The theory of the second variation has been elaborately developed by many writers. The most important of the early papers is that of Clebsch [29] in which he transformed the second variation into its so-called reduced form and derived therefrom his necessary condition analogous to that of Legendre for simpler problems. The methods of Clebsch were modified by A. Mayer [30] who proved the necessity of a condition analogous to that of Jacobi for simpler problems, the so-called condition of Mayer described in the preceding pages. In a series of papers von Escherich [31] discussed in great detail the theory of the second variation and the various consequences which can be deduced from it. A condensed treatment of his theory is given by Bolza [32]. Hahn [33] showed the relationship between the theory of the second variation and certain aspects of the theories of Weierstrass as extended to the problem of Lagrange. The theory of the second variation takes a relatively simple form when it is viewed from the stand-point of the theory of the minimum problem of the second variation, as has been shown by Bliss [27, 34, 35].

The best reference for the sufficiency theorems in Chapter IV above is Bolza [36] to whom the precise formulation of the theorems and many details of the proofs are due. The properties of fields and their relation to the invariant integral analogous to that of Hilbert for simpler cases were first discussed by A. Mayer [37], and further material pertinent to the sufficiency proofs was discussed by Bolza [38] and Carathéodory [39]. The reader may refer to Kneser [11, 2d ed., pp. 290 ff.] for sufficiency proofs for the Mayer problem, and to Bliss [35] for a proof of the integral formula of Weierstrass and other properties of fields for the Lagrange problem.

BIBLIOGRAPHY.

GENERAL REFERENCES.

1. Kneser, "Variationsrechnung," *Encyclopädie der Mathematischen Wissenschaften*, II A 8; Zermelo und Hahn, "Weiterentwicklungen der Variationsrechnung in den letzten Jahren," *ibid.*, II A 8a.

2. Lecat, "Calcul des Variations," *Encyclopédie des Sciences Mathématiques*, II 31 (1913 and 1916).

3. Bolza, *Vorlesungen über Variationsrechnung* (1909).

4. Hadamard, *Leçons sur le Calcul des Variations* (1910).

5. Bliss, *The Calculus of Variations* (1924).

6. Moigno-Lindelöf, *Calcul des Variations* (1861).

THE EULER-LAGRANGE MULTIPLIER RULE.

7. Euler, *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes* (1744).

8. Lagrange, *Oeuvres* I, X. See also the translation in Ostwald's *Klassiker der Exakten Wissenschaften*, Nr. 47, pp. 31 ff.

9. A. Mayer, "Begründung der Lagrangeschen Multiplicatorenmethode der Variationsrechnung," *Mathematische Annalen*, Vol. 26 (1886), pp. 74-82.

10. Hilbert, "Zur Variationsrechnung," *Göttinger Nachrichten* (1905), pp. 159-180; *Mathematische Annalen*, Vol. 62 (1906), pp. 351-368.

11. Kneser, *Lehrbuch der Variationsrechnung* (1900); 2d ed. (1925).

12. Hahn, "Über die Lagrangeschen Multiplicatorenmethode in der Variationsrechnung," *Monatshefte für Mathematik und Physik*, Vol. 14 (1908), pp. 325-342.

13. Hahn, "Über die Herleitung der Differentialgleichungen der Variationsrechnung," *Mathematische Annalen*, Vol. 63 (1907), pp. 253-272.

14. Bliss, "The Solutions of Differential Equations of the First Order as Functions of Their Initial Values," *Annals of Mathematics*, 2d Series, Vol. 6 (1905), pp. 49-68.

15. Bliss, "Solutions of Differential Equations as Functions of the Constants of Integration," *Bulletin of the American Mathematical Society*, Vol. 24 (1918), pp. 15-26.

16. Bliss, "The Problem of Mayer with Variable End-Points," *Transactions of the American Mathematical Society*, Vol. 19 (1918), pp. 305-314.

NORMAL ARCS.

17. von Escherich, see reference 31 below, Vol. 108 (1899), p. 1290.

18. Hahn, "Bemerkungen zur Variationsrechnung," *Mathematische Annalen*, Vol. 58 (1904), pp. 148-168.

19. Bolza, "Über den 'Anormalen Fall' beim Lagrangeschen und Mayerschen Problem mit gemischter Bedingungen und variabeln Endpunkten," *Mathematische Annalen*, Vol. 74 (1913), pp. 430-436.

THE NECESSARY CONDITIONS OF WEIERSTRASS AND CLEBSCH.

20. Clebsch, "Über die Reduction der zweiten Variation auf ihre einfachste Form," *Journal für die reine und angewandte Mathematik*, Vol. 55 (1858), pp. 254-273.

21. Hahn, "Über das allgemeine Problem der Variationsrechnung," *Monatshefte für Mathematik und Physik*, Vol. 17 (1906), pp. 295-304.

22. Bolza, see reference 3 above, pp. 603 ff.

23. Goursat, *Cours d'analyse*, Vol. 3, 4th ed. (1927), pp. 545-660.

THE NECESSARY CONDITION OF MAYER.

24. A. Mayer, "Über die Kriterien des Maximums und Minimums der einfachen Integrale," *Journal für die reine und angewandte Mathematik*, Vol. 69 (1868), pp. 238-263.

25. Kneser, "Die Jacobische Bedingung des Extremums bei einem allgemeinen Typus von Aufgaben der Variationsrechnung," *Mittheilungen der mathematischen Gesselschaft zu Kharkov*, Ser. 2, Vol. 7 (1902), pp. 253-267.

26. Bolza, see reference 3 above, pp. 613 ff.

27. Bliss, "Jacobi's Condition for Problems of the Calculus of Variations in Parametric Form," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 195-206.

28. D. M. Smith, "Jacobi's Condition for the Problem of Lagrange in the Calculus of Variations," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 459-475.

THE THEORY OF THE SECOND VARIATION.

29. Clebsch, see reference 20 above.

30. A. Mayer, see reference 24 above.

31. von Escherich, "Die zweite Variation der einfachen Integrale," *Sitzungsberichten der kaiserlichen Akademie der Wissenschaften in Wien*, Vol. 107 (1898), pp. 1191-1250, 1267-1326, 1383-1430; Vol. 108 (1899), pp. 1269-1340; Vol. 110 (1901), pp. 1355-1421.

32. Bolza, see reference 3 above, pp. 619-634.

33. Hahn, "Über den Zusammenhang zwischen den Theorien der zweiten Variation und der Weierstrass'schen Theorie die Variationsrechnung," *Rendiconti del Circolo Matematico di Palermo*, Vol. 29 (1910), pp. 49-78.

34. Bliss, "A Note on the Problem of Lagrange in the Calculus of Variations," *Bulletin of the American Mathematical Society*, Vol. 22 (1916), pp. 220-225.

35. Bliss, "The Transformation of Clebsch in the Calculus of Variations," *Proceedings of the International Mathematical Congress Held in Toronto*, Vol. 1 (1924), pp. 589-603.

SUFFICIENT CONDITIONS FOR A MINIMUM.

36. Bolza, see reference 3, pp. 635 ff.

37. A. Mayer, "Über den Hilbertschen Unabhangigkeitssatz in der Theorie des Maximums und Minimums der einfacher Integrale," *Leipziger Berichte*, Vol. 55 (1903), pp. 131-145.

38. Bolza, "Weierstrass' Theorem and Kneser's Theorem on Transversals for the Most General Case of an Extremum of simple Definite Integrals," *Transactions of the American Mathematical Society*, Vol. 7 (1906), pp. 459-488.

39. Carathéodory, "Die Methode der Geodätische Equidistanten und das Problem von Lagrange," *Acta Mathematica*, Vol. 47 (1926), pp. 199-236.

40. Kneser, see the reference 11, 2d ed., pp. 290 ff.

41. Bliss, see reference 35, p. 593.

Finite Geometries and the Theory of Groups.*

By R. D. Carmichael.

Introduction.

The general purpose of this memoir is to exhibit the close contact which exists between the finite projective geometries $PG(k, p^n)$ and the theory of finite groups and to utilize the geometry in constructing permutation groups and in investigating their properties. Special attention is given to the case of multiply transitive permutation groups.

In the first division (§§ 1-6) a representation is given of the finite projective geometries $PG(k, p^n)$ by means of Abelian groups of type $(1, 1, 1, \cdots)$ and order $p^{(k+1)n}$ where p is prime. For the purpose of effecting this representation a system of coördinates for denoting the elements of such an Abelian group is introduced by means of the marks of the Galois field $GF[p^n]$. It is believed that these coördinates will be found useful for other purposes than those to which they are here put. They are used to aid in the selection and definition of a normal set of subgroups, which subgroups are interpreted as the points of a finite projective geometry $PG(k, p^n)$ of k dimensions. The elements themselves of the given Abelian group then become the points of a Euclidean geometry $EG(k+1, p^n)$ of $k+1$ dimensions. The theory of the finite geometries thus becomes available for developing the theory of Abelian groups of type $(1, 1, 1, \cdots)$, and *vice versa.* In particular, it is shown in § 6 that every theorem relating to a general projective space or a proper projective space or a modular projective space or a rational modular projective space (in the sense of Veblen and Young, l. c.) may be translated into a theorem about Abelian groups of type $(1, 1, 1, \cdots)$. Thus by a single act of thought a significant extension is given to the theory of Abelian groups and a method is made apparent by which the theory may be further developed.

By means of the coördinates introduced in § 2 to denote the elements of an Abelian group G of order $p^{(k+1)n}$ and of type $(1, 1, 1, \cdots)$ analytical representations are set up in the second division of the memoir (§§ 7-11) for the group of isomorphisms of the named Abelian group and for its holomorph. These representations afford generalisations of known results. Incidentally to the study of certain subgroups of the group of isomorphisms a generalisation of the Betti-Mathieu group appears (§ 9). Finally, in the

* Presented to the American Mathematical Society (Kansas City), Dec. 29, 1925.

last section of this division certain transformation groups in $PG(k+1, p^n)$ are formed from the earlier groups in the division by aid of the interpretation of the given Abelian group by means of the Euclidean space $EG(k+1, p^n)$.

The third division of the memoir (§§ 12-14) is devoted to the development of certain central theorems concerning collineation groups in the finite geometries and their subgroups. The main results are given in the first and third theorems of § 12 and the first theorem of § 13. The results include and generalise several known theorems concerning doubly transitive and triply transitive groups. In particular the existence is shown of several infinite classes of triply transitive and of doubly transitive groups, including certain such classes already known. Moreover, infinite classes of simply transitive primitive groups are also exhibited. It is proved that there is no upper limit K to the number of primitive groups (of varying degrees) in a set of primitive groups each of which is simply isomorphic with each of the others in the set. Furthermore, it is shown that, for every integer L there exist integers $s[t]$ such that the number of the doubly transitive [triply transitive] groups of degree $s[t]$ is greater than L.

I. REPRESENTATION OF THE FINITE PROJECTIVE GEOMETRIES $PG(k, p^n)$ BY MEANS OF ABELIAN GROUPS.

1. *The Finite Projective Geometry $PG(k, p^n)$.* Let p be any prime number and k be any positive integer. Let us consider the Abelian group G_{k+1} of order p^{k+1} and type $(1, 1, 1, \cdots)$. Every element of this group except the identity is of order p. The number of these elements is $p^{k+1}-1$. Each of them generates a subgroup of order p, the same subgroup being generated by each of $p-1$ different elements. Hence the group G_{k+1} contains

$$(p^{k+1}-1)/(p-1), \text{ or } 1+p+p^2+\cdots+p^k,$$

distinct subgroups of order p. The totality of these subgroups contains all the elements of G_{k+1}; and no two of these subgroups have any element in common except identity.

Each of these subgroups of order p in G_{k+1} will be called a point in the finite geometry $PG(k, p)$ which we are engaged in constructing. This k-dimensional finite geometry then contains just $1+p+p^2+\cdots+p^k$ points.

Now consider any two points of the $PG(k, p)$. From the group-theoretic point of view they are two subgroups of G_{k+1} of order p. The group generated by them is of order p^2 and type $(1, 1)$. It contains $1+p$ subgroups of order p; and no two of these subgroups have any element in common except

identity. From the geometric point of view these $1+p$ subgroups are $1+p$ points of the $PG(k,p)$. We shall say that they form a line in the $PG(k,p)$. Thus any two points in $PG(k,p)$ determine a line of $PG(k,p)$, and this line has just $1+p$ points on it. We shall denote by AB the line containing the two distinct points A and B.

Let us determine the number of lines in $PG(k,p)$. In determining a line we may select a first point in $1+p+p^2+\cdots+p^k$ ways and then a second point in $p+p^2+\cdots+p^k$ ways. But this procedure will select the same line in as many ways as two points on it may be chosen in an assigned order. The first point may be taken in $1+p$ ways and then the second in p ways. Hence the number of lines in $PG(k,p)$ is

$$(1+p+p^2+\cdots+p^k)(p+p^2+\cdots p^k)/(1+p)p$$

or

$$(p^{k+1}-1)(p^k-1)/(p^2-1)(p-1).$$

This of course is the same as the number of subgroups of order p^2 in G_{k+1}.

An m-dimensional space in $PG(k,p)$, $m\leqq k$, may now be defined as the set of points each of which is identified with the corresponding subgroup of order p in a given subgroup of G_{k+1} of order p^{m+1}, its type of course being necessarily $(1,1,1,\cdots)$. For $m=2$ we have the case of a plane. The number of points in the m-dimensional space is

$$1+p+p^2+\cdots+p^m.$$

It is obvious that this m-dimensional space is completely determined by any $m+1$ of its points so selected that they do not all lie in any $(m-1)$-dimensional space.

In this m-dimensional space $PG(m,p)$ there are included $(m-1)$-dimensional spaces $PG(m-1,p)$. Let us consider any such space S_{m-1} of $m-1$ dimensions, m now being greater than 1; and let P be a point not in S_{m-1}. Let T be the set of points each of which is collinear (on the same line) with P and some point of S_{m-1}. From the group-theoretic interpretation it is clear that the set of points T constitute an m-dimensional space $PG(m,p)$. Thus we may have an inductive definition of the points of a space of m dimensions. A point is a 0-space. If $P_1, P_2, \cdots, P_{m+1}$ are points not all in the same $(m-1)$-space, then the set of all points each of which is collinear with P_{m+1} and some point of the $(m-1)$-space $(P_1, P_2, \cdots, P_m)$ is the m-space $(P_1, P_2, \cdots, P_{m+1})$. It is obvious that this inductive definition is equivalent to the definition already given.

The number of ways in which $m+1$ points may be selected in a given order so that they do not all lie in any $(m-1)$-dimensional space is

$$(1+p+p^2+\cdots+p^k)(p+p^2+\cdots p^k)$$
$$\times(p^2+p^3+\cdots+p^k)\cdots(p^m+p^{m+1}+\cdots+p^k),$$

the factors of this expression in the order written being the number of ways in which the first point, the second point, $\cdots$, the $(m+1)$-th point, respectively, may be selected. The number of ways in which $m+1$ points of a given $PG(m,p)$ may be selected in a given order so that they do not all lie on any $(m-1)$-dimensional space is

$$(1+p+\cdots+p^m)(p+p^2+\cdots+p^m)$$
$$\times(p^2+p^3+p^m)\cdots(p^{m-1}+p^m)p^m.$$

It is obvious that the number of m-dimensional spaces $PG(m,p)$ in the given $PG(k,p)$ is the quotient of the first of the two foregoing products divided by the second; this quotient may be written in the form

$$\frac{(p^{k+1}-1)(p^k-1)(p^{k-1}-1)\cdots(p^{k-m+1}-1)}{(p^{m+1}-1)(p^m-1)(p^{m-1}-1)\cdots(p^2-1)(p-1)}.$$

This of course is the same as the number of subgroups in G_{k+1} of order p^{m+1}.

When $m+1$ generators of G_{k+1} are selected for generating the subgroups corresponding to a given $PG(m,p)$ there are left in G_{k+1} $k-m$ other independent generators independent of the $m+1$ already employed. These give rise to a $PG(k-m-1,p)$. Thence we see that the number of m-spaces in $PG(k,p)$ is the same as the number of $(k-m-1)$-spaces. In particular, the number of points in a plane is equal to the number of lines in the plane.

Veblen and Bussey * define a finite projective geometry in the following way. It consists of a set of elements, called points for suggestiveness, which are subject to the following five conditions or postulates:

I. The set contains a finite number (>2) of points. It contains one or more subsets called lines, each of which contains at least three points.

II. If A and B are distinct points, there is one and only one line that contains both A and B.

III. If A, B, C are non-collinear points and if a line l contains a point D of the line AB and a point E of the line BC but does not contain A or B or C, then the line l contains a point F of the line CA.

IV_k. If m is an integer less than k, not all of the points considered are in the same m-space.

V_k. If IV_k is satisfied, there exists in the set of points considered no $(k+1)$-space.

---

* O. Veblen and W. H. Bussey, "Finite Projective Geometrics," *Transactions of the American Mathematical Society*, Vol. 7 (1906), pp. 241-259.

The geometry so defined is a geometry of k-dimensional space.

In this system of postulates the terms point and line are left undefined. A point is called a 0-space and a line is called a 1-space. Spaces of higher dimensions are defined inductively by the method which we have already shown to be equivalent to our first definition of an m-space. To show that the set of points which we have defined constitute a finite projective geometry it is therefore sufficient to prove that each of the foregoing postulates is satisfied. From the properties of the group G_{k+1} it follows at once that postulates I, II, IV_k, V_k are satisfied by the set of points in $PG(k,p)$. It remains to show that postulate III is verified. For this purpose let a, b, c be generators of the subgroups of G_{k+1} corresponding to the points A, B, C respectively. Then the groups corresponding to the points of the lines AB, BC, CA have respectively as generators the elements

$$a^{\alpha_1}b^{\beta_1}, \quad b^{\beta_2}c^{\alpha_2}, \quad c^{\gamma_3}a^{\alpha_3},$$

where each exponent belongs to the set $0, 1, 2, \cdots, p-1$ and at least one exponent in the symbol for each generator is different from zero. If a generator of the group corresponding to D in the postulate is $a^{\alpha}b^{\beta}$ then both α and β belong to the set $1, 2, \cdots, p-1$, since D is different from A and B. Likewise both ρ and σ in a generator $b^{\rho}c^{\sigma}$ of the group corresponding to E belong to the set $1, 2, \cdots, p-1$. Then the line DE corresponds to the group $\{a^{\alpha}b^{\beta}, b^{\rho}c^{\sigma}\}$. The elements in this group are $a^{\lambda\alpha}\, b^{\lambda\beta}\, b^{\mu\rho}\, c^{\mu\sigma}$ where λ and μ range independently over the set $0, 1, 2, \cdots, p-1$. Now λ and μ, both different form zero, exist such that $\lambda\beta + \mu\rho \equiv 0$ modulo p. The corresponding element of the group is then $a^{\lambda\alpha}c^{\mu\sigma}$. This generates a group corresponding to a point on the line AC; it is different from A and C since each of the numbers α, σ, λ, μ is incongruent to zero modulo p. This is the point F common to DE and CA whose existence is asserted by postulate III. Hence the set of points in our $PG(k,p)$ satisfies the foregoing postulates and therefore constitutes a finite projective geometry.*

It is desirable to introduce homogeneous coördinates for representing the points in the finite projective geometry $PG(k,p)$. For this purpose let us consider a set of $k+1$ independent generators $a_0, a_1, a_2, \cdots, a_k$ of the group G_{k+1}. Then the elements of this group are all represented uniquely by the set of symbols

$$a_0^{\mu_0}a_1^{\mu_1}a_2^{\mu_2}\cdots a_k^{\mu_k}$$

* The special case of the geometry $PG(3,2)$ is treated briefly in a manner similar to the foregoing by U. G. Mitchell in his dissertation (footnote on p. 34).

where $\mu_0, \mu_1, \cdots, \mu_k$ run independently over the set $0, 1, 2, \cdots, p-1$ of p numbers. An element of G_{k+1} may therefore be denoted uniquely by the symbol

$$\{\mu_0, \mu_1, \mu_2, \cdots, \mu_k\}$$

where each μ is a number of the set $0, 1, 2, \cdots, p-1$, provided it is understood that the symbol represents the product $a_0^{\mu_0}a_1^{\mu_1} \cdots a_k^{\mu_k}$, the a's forming a fixed set of independent generators of G_{k+1}. Two such symbols are to be considered equivalent if their corresponding elements are congruent modulo p. For the multiplication of these symbols (corresponding to multiplication of elements in G_{k+1}) we obviously have the following formula

$$\{\mu_0, \mu_1, \cdots, \mu_k\}\{\nu_0, \nu_1, \cdots, \nu_k\} = \{\mu_0 + \nu_0, \mu_1 + \nu_1, \cdots, \mu_k + \nu_k\}.$$

Now consider the set of elements

$$\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$$

where $\mu_0, \mu_1, \cdots, \mu_k$ constitute a fixed set of $k+1$ numbers taken modulo p and not all of them are congruent to zero modulo p, μ being a variable integer taken modulo p. It is easy to see that this set of elements forms a group of order p having $\{\mu_0, \mu_1, \cdots, \mu_k\}$ for a generator. This group may be denoted by the symbol

$$(\mu_0, \mu_1, \cdots, \mu_k).$$

The same group is also represented by the symbol

$$(\rho\mu_0, \rho\mu_1, \cdots, \rho\mu_k)$$

provided only that ρ is a fixed integer incongruent to zero modulo p. The corresponding point will be denoted by the symbol

$$(\mu_0, \mu_1, \cdots, \mu_k)$$

and $\mu_0, \mu_1, \cdots, \mu_k$ will be called homogeneous coördinates of the point. The condition that such a symbol shall represent a point is that the μ's shall be integers and that one of them at least shall be different from zero modulo p. Two such symbols represent the same point if the corresponding coördinates are proportional modulo p. Except for this factor of proportionality there is thus a unique correspondence between the points of $PG(k, p)$ and the symbols which represent them by means of coördinates.

2. *Generalization to the Finite Projective Geometry $PG(k, p^n)$.* Let us consider more generally an Abelian group $G_{(k+1)n}$ of order $p^{(k+1)n}$ and type $(1, 1, 1, \cdots)$, p being a prime number and k and n being any positive in-

tegers. The points of our finite geometry $PG(k, p^n)$ are to be certain subgroups of $G_{(k+1)n}$ of order p^n. To begin with, these subgroups* are to be selected in such a way † that no two of them shall have any element in common except identity and so that the set shall contain all the elements of $G_{(k+1)n}$. The number of elements other than identity in $G_{(k+1)n}$ is $p^{(k+1)n}-1$; and the number of such elements in a subgroup of order p^n is p^n-1. Hence a set of subgroups of $G_{(k+1)n}$ of order p^n and having the properties named will consist of

$$(p^{(k+1)n}-1)/(p^n-1), \text{ or } 1+p^n+p^{2n}+\cdots+p^{kn},$$

subgroups. Therefore the k-dimensional geometry $PG(k, p^n)$, to be defined, will consist of $1+p^n+p^{2n}+\cdots+p^{kn}$ points.

In order to select an appropriate set of subgroups of order p^n for the purpose in hand we shall first develop a method of representing the elements of $G_{(k+1)n}$ by means of the marks of a Galois field, thus generalizing the results at the end of the preceding section. This mode of representing the elements of an Abelian group of type $(1, 1, 1, \cdots)$ we shall find useful for other purposes besides the geometrical one which now engages our attention.

Let us denote a set of $(k+1)n$ independent generating elements of $G_{(k+1)n}$ by

$$\begin{array}{ccccc} a_{01}, & a_{02}, & a_{03}, \cdots, & a_{0n}, \\ a_{11}, & a_{12}, & a_{13}, \cdots, & a_{1n}, \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1}, & a_{k2}, & a_{k3}, \cdots, & a_{kn}. \end{array}$$

Then every element in $G_{(k+1)n}$ may be represented uniquely in the form

$$\prod_{i=0}^{k} a_{i1}^{s_{i1}} a_{i2}^{s_{i2}} \cdots a_{in}^{s_{in}}$$

where the exponents s are integers taken modulo p. The element denoted by this product for a fixed set of exponents s will be represented by the symbol

$$\{\mu_0, \mu_1, \mu_2, \cdots, \mu_k\}$$

where μ_i $(i=0, 1, 2, \cdots, k)$ denotes that mark of the Galois field $GF[p^n]$ which may be written in the form

$$\mu_i = s_{i1} + s_{i2}\omega + s_{i3}\omega^2 + \cdots + s_{in}\omega^{n-1},$$

* The special case when $p=2$ and $k=1$ is treated incidentally (in a different manner) by L. E. Dickson, *Bulletin of the American Mathematical Society*, Ser. 2, Vol. 11 (1905), pp. 177-179.

† See G. A. Miller, *Bulletin of the American Mathematical Society*, Ser. 2, Vol. 12 (1906), pp. 446-449, for theorems relating to this problem.

ω being a fixed primitive mark of $GF[p^n]$. This correspondence of elements and symbols is unique in the sense that to each element there corresponds a single symbol and to each symbol there corresponds a single element.

For the multiplication of these symbols, corresponding to the multiplication of elements in $G_{(k+1)n}$, we have the following obvious formula:

$$\{\mu_0, \mu_1, \cdots, \mu_k\}\{\nu_0, \nu_1, \cdots, \nu_k\} = \{\mu_0 + \nu_0, \mu_1 + \nu_1, \cdots, \mu_k + \nu_k\}.$$

Now suppose that $\mu_0, \mu_1, \cdots, \mu_k$ is a fixed set of $k+1$ marks of $GF[p^n]$, at least one of them being different from zero; and consider the set of elements

$$\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$$

where μ is a variable running over the p^n marks of $GF[p^n]$. It is obvious that the elements in this set are all distinct and that their number is p^n. Moreover, the product of any two of them is in the set, as one sees immediately from the law of multiplication and the properties of the marks of a Galois field. This set of elements therefore constitutes a subgroup of $G_{(k+1)n}$ of order p^n. It is easy to see that the elements

$$\{\omega^i\mu_0, \omega^i\mu_1, \cdots, \omega^i\mu_k\}, \qquad (i = 0, 1, 2, \cdots, n-1),$$

constitute a set of independent generators of this subgroup. If σ is any non-zero mark of the Galois field the same subgroup obviously consists of the set of elements

$$\{\mu\sigma\mu_0, \mu\sigma\mu_1, \cdots, \mu\sigma\mu_k\},$$

μ varifying as before. The subgroup itself may therefore be represented by the symbol

$$(\mu_0, \mu_1, \cdots, \mu_k)$$

where $\mu_0, \mu_1, \cdots, \mu_k$ are interpreted as the "homogeneous coördinates" of the subgroup. On multiplying each of the coördinates by one and the same non-zero mark of the field we have merely proportional homogeneous coördinates of the same subgroup. To each set of ordered coördinates, one at least of the coördinates being different from zero, there corresponds a subgroup of $G_{(k+1)n}$ of order p^n.

The number of subgroups in the set denoted by $(\mu_0, \mu_1, \cdots, \mu_k)$ for varying μ's is readily determined. Each symbol μ may be chosen in p^n independent ways except that they cannot all be zero. Hence the number of choices is $p^{(k+1)n} - 1$. To obtain the number of subgroups we must divide

this by the number p^n-1 of possible factors of proportionality in the various notations for the same group. Hence the number of groups in our set is

$$(p^{(k+1)n}-1)/(p^n-1), \text{ or } 1+p^n+p^{2n}+\cdots+p^{kn}.$$

Once $G_{(k+1)n}$ has been given, this selection of groups depends on two things: the ordered set of $(k+1)n$ independent generators and the primitive mark ω by means of which the marks μ_i were first introduced. With reference to this selected basis of determination we shall call the set of subgroups just determined a normal set. By means of other sets of generators and other primitive marks we might in certain cases select other normal sets of subgroups of $G_{(k+1)n}$. Since we shall use the same basis throughout this memoir we shall speak of the foregoing normal set of subgroups without reference to the basis on which it has been defined.

For the case $n=1$, it is to be observed, a subgroup of a normal set is simply any subgroup of order p.

No two subgroups of a normal set have any element in common except identity, as one may readily prove by means of the symbols which represent their elements. Moreover a given element of $G_{(k+1)n}$ occurs in some subgroup of a normal set. Hence the subgroups of a normal set have the properties demanded at the beginning of the section for points. Accordingly for the points of $PG(k, p^n)$ we take the subgroups of a normal set of subgroups of $G_{(k+1)n}$. That the latter group has (when $n>1$) other subgroups of the same order p^n will not concern us at the present.

We shall represent a point of $PG(k, p^n)$ by the same symbol

$$(\mu_0, \mu_1, \cdots, \mu_k)$$

as we have already employed to denote the subgroup of order p^n which we identify with this point. Thus we have a set of homogeneous coördinates to represent the points of $PG(k, p^n)$, each one of the coördinates being a mark of $GF[p^n]$.

An m-dimensional space, or an m-space, in $PG(k, p^n)$, $m \leqq k$, may now be defined as the set of points corresponding to the groups of a normal set of subgroups of $G_{(k+1)n}$ which are contained as subgroups in the group generated by $m+1$ of the groups of a normal set, these $m+1$ groups being such that no one of them is contained in the group generated by the other m. A point will be called a 0-space; a 1-space will be called a line; a 2-space we will call a plane. It is clear that this definition is again equivalent to the inductive definition given in § 1 for the special case when $n=1$.

To show that the $PG(k, p^n)$ is a finite projective geometry in the sense

of Veblen and Bussey we have now only to prove that the postulates given in § 1 are verified when interpreted as referring to our $PG(k, p^n)$. That postulates I, II, IV_k, V_k hold is immediately obvious. It remains only to verify postulate III.

For this purpose consider three non-collinear points A, B, C and let

$$(\alpha_0, \alpha_1, \cdot\cdot\cdot, \alpha_k), \quad (\beta_0, \beta_1, \cdot\cdot\cdot, \beta_k), \quad (\gamma_0, \gamma_1, \cdot\cdot\cdot, \gamma_k)$$

respectively be their coördinates. Then a point D on the line AB determined by the points A and B has the coördinates

$$(\alpha\alpha_0 + \beta\beta_0, \ \alpha\alpha_1 + \beta\beta_1, \cdot\cdot\cdot, \alpha\alpha_k + \beta\beta_k)$$

where α and β are marks of $GF[p^n]$. A necessary and sufficient condition that this point shall be different from A and B is that both α and β shall be different from zero. Hence we take them to be different from zero. Likewise a point E on BC has the coördinates

$$(\rho\beta_0 + \sigma\gamma_0, \ \rho\beta_1 + \sigma\gamma_1, \cdot\cdot\cdot, \rho\beta_k + \sigma\gamma_k)$$

where ρ and σ are marks of $GF[p^n]$. We take ρ and σ to be both different from zero so that E shall be different from both B and C. Now a point on the line DE has the coördinates

$$(\lambda\alpha\alpha_0 + \lambda\beta\beta_0 + \mu\rho\beta_0 + \mu\sigma\gamma_0, \cdot\cdot\cdot, \lambda\alpha\alpha_k + \lambda\beta\beta_k + \mu\rho\beta_k + \mu\sigma\gamma_k)$$

where λ and μ are marks of $GF[p^n]$. Since β and ρ are both different from zero there exist non-zero marks λ and μ such that $\lambda\beta + \mu\rho$ is zero. For such a pair of values of λ and μ the corresponding point F of DE has the coördinates

$$(\lambda\alpha\alpha_0 + \mu\sigma\gamma_0, \cdot\cdot\cdot, \lambda\alpha\alpha_k + \mu\sigma\gamma_k).$$

This F is a point on the line CA; and it is different from both C and A, since each of the marks α, β, λ, μ is different from zero. From the relation thus established among the points A, B, C, D, E, F it is seen that postulate III is verified.

Hence, the $PG(k, p^n)$, as we have defined it, is a finite projective geometry in the sense of Veblen and Bussey.

Veblen and Bussey (*loc. cit.*) proved that when $k > 2$ every finite projective k-dimensional geometry satisfying the definition which we have reproduced in § 1 is a geometry of points whose homogeneous coördinates may be taken as the marks of $GF[p^n]$ in precisely the same way as we have used homogeneous coördinates to represent the points of our $PG(k, p^n)$. This justifies us in using for these geometries the symbol $PG(k, p^n)$ already em-

ployed by Veblen and Bussey. Moreover, we may say that the foregoing group-theoretic construction of $PG(k, p^n)$ affords an interpretation in the theory of Abelian groups of type $(1, 1, 1, \cdots)$ of every possible finite projective geometry of more than two dimensions. We shall not now treat the problem of possible group-theoretic interpretations of the remaining finite geometries, namely, certain of those of two dimensions.

It is now evident that every theorem relating to $PG(k, p^n)$ can be translated into a corresponding theorem about the group $G_{(k+1)n}$. We shall illustrate the remark by so interpreting the following geometric theorem:

If l and m are positive integers less than k and such that $l + m - k = r \geqq 0$, then, in the given k-space, an l-space and an m-space have at least an r-space in common.

The group-theoretic interpretation is as follows:

Let $s_1, s_2, \cdots, s_{l+1}$ be any $l+1$ subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is contained in the group generated by the other l, and let $\sigma_1, \sigma_2, \cdots, \sigma_{m+1}$ be a like set of $m+1$ such subgroups. Then, if $l < k$, $m < k$, $l + m - k = r \geqq 0$, the groups $\{s_1, s_2, \cdots, s_{l+1}\}$ and $\{\sigma_1, \sigma_2, \cdots, \sigma_{m+1}\}$ contain at least $r+1$ subgroups of a normal set such that no one of these subgroups is contained in the group generated by the remaining r of them.

The number of m-spaces $PG(m, p^n)$, $m < k$, contained in the given k-space $PG(k, p^n)$ is readily determined in the general case by the same method as that employed in § 1 for the special case $n = 1$. This number turns out to be

$$\frac{(p^{(k+1)n} - 1)(p^{kn} - 1)(p^{(k-1)n} - 1) \cdots (p^{(k-m+1)n} - 1)}{(p^{(m+1)n} - 1)(p^{mn} - 1)(p^{(m-1)n} - 1) \cdots (p^n - 1)}.$$

In the foregoing part of the section we have given an analytic method for determining normal sets of subgroups of $G_{(k+1)n}$. It is desirable to have such a set characterised by means of properties which are immediately group-theoretic in their character. The subgroups of a given normal set have the following properties and mutual relations, as we have already seen:

1) Each of these subgroups is of order p^n.
2) No two of these subgroups have a common element except identity.
3) Any given element of $G_{(k+1)n}$ is contained in some subgroup of a normal set.
4) If A, B, C are three subgroups of a normal set such that no one of

them is in the group generated by the other two, and if D is a subgroup of the group $\{A, B\}$ and is different from A and B and belongs to the normal set, and finally if E is a subgroup of the group $\{B, C\}$ and is different from B and C and belongs to the normal set, then the groups $\{C, A\}$ and $\{D, E\}$ have in common a group F which belongs to the normal set.

Now any set of subgroups of $G_{(k+1)n}$ which have these properties alone clearly satisfy the five defining postulates given in § 1. They therefore afford a representation of a finite geometry. But Veblen and Bussey (*loc. cit.*) have shown that every finite projective k-dimensional geometry satisfying the definition reproduced in § 1 is a $PG(k, p^n)$, in the sense of their use of this symbol, provided that $k > 2$. Hence one can introduce coördinates into this geometry by means of the marks of a Galois field. On doing this in the case of the given group-theoretic representation of the geometry we exhibit the set of subgroups involved as a normal set in accordance with the definition of such a set. Therefore when $k > 2$ the properties 1), 2), 3), 4) of a normal set of subgroups furnish a complete group-theoretic characterization of such a set. The conclusion will also hold for $k = 1$ or 2 if we suppose that the normal set of subgroups is so chosen that it may be taken as a part of the normal set of subgroups in a group of order p^{4n} and type $(1, 1, 1, \cdots)$ which contains the given group $G_{(k+1)n}$ for $k = 1$ or 2.

Consider now the $PG(k+1, p^n)$ whose points are denoted by the symbols $(\mu_0, \mu_1, \cdots, \mu_{k+1})$ where each μ is a mark of $GF[p^n]$. Those points for which μ_{k+1} is different from zero constitute the Euclidean finite geometry $EG(k+1, p^n)$, this being obtained by omitting from $PG(k+1, p^n)$ those points for which the last coördinate is zero. For the points of $EG(k+1, p^n)$ we may take $\mu_{k+1} = 1$. Then the coördinates $\mu_0, \mu_1, \cdots, \mu_k$ may be taken as the non-homogeneous coördinates of points in $EG(k+1, p^n)$. Such a point $(\mu_0, \mu_1, \cdots, \mu_k, 1)$ may then be identified with the element $\{\mu_0, \mu_1, \cdots, \mu_k\}$ of $G_{(k+1)n}$. Hence the elements of this group may be taken as the points of the Euclidean finite geometry $EG(k+1, p^n)$. Hence the theorems in the latter geometry may be interpreted as theorems concerning the elements of $G_{(k+1)n}$.

3. *The Principle of Duality.* The principle of duality is valid in the finite geometry $PG(k, p^n)$. If l is less than k the dual of the set of l-spaces in $PG(k, p^n)$ is the set of $(k-l-1)$-spaces. In particular the dual of the set of points in $PG(k, p^n)$ is the set of $(k-1)$-spaces contained in the given k-space. Since the number of elements [l-spaces] in a set of subspaces is

equal to the number of elements [$(k-l-1)$-spaces] in the dual set of spaces, it follows in particular that the number of subgroups of a normal set of subgroups of $G_{(k+1)n}$ is equal to the number of subgroups each of which is generated by $k-1$ independent subgroups of the normal set. For $n=1$ this reduces to the well known theorem that the number of subgroups of order p in an Abelian group of order p^{k+1} and type $(1, 1, 1, \cdots)$ is equal to the number of subgroups of index p. More generally, if l is less than $k+1$ the number of subgroups of order p^l in this group G_{k+1} is equal to the number of subgroups of index p^l.

In general every theorem about the Abelian group of order p^m and type $(1, 1, 1, \cdots)$, which is capable of interpretation as a theorem in a finite geometry $PG(k, p^n)$, may be dualized. It will thus lead to a new theorem about the Abelian group, except in the special case when the theorem is its own dual. For the purpose of obtaining these theorems about a given Abelian group of order p^m and type $(1, 1, 1, \cdots)$, one may construct a corresponding geometry $PG(k, p^n)$ for every pair of positive integral values of k and n such that $(k+1)n=m$. Thus if m is highly composite the given Abelian group may be investigated by means of any one of several finite geometries constructed in the manner indicated. The case $k+1=m$ and $n=1$ will be especially useful for this purpose since in this case the normal set of subgroups consists of all the subgroups of order p.

From the principle of duality it follows that one of the requirements for points named at the beginning of § 2 is superfluous, at least in the form there stated. It was prescribed that the subgroups which were to represent points were to be selected in such a way that no two of them should have any element in common except identity. Now that the geometry has been constructed a new one can be made from it such that the points in the new geometry are the dual elements of the points in the old geometry. In this new geometry two given points, when considered as subgroups, will have elements in common besides the identity. And yet the new geometry will serve equally well as a means of investigating the given Abelian group.

Once this general principle of duality in the theory of Abelian groups is recognized, a number of properties of these groups heretofore discovered become almost or quite obvious, since a fundamental reason for their appearance is manifest.

4. *The Complete Quadrangle.* In the finite projective geometries $PG(k, p^n)$ there is an important distinction to be made according as the prime p is equal to 2 or is odd. This distinction was investigated by Veblen

and Bussey in the article cited. They showed (p. 245) that the diagonal points of a complete quadrangle are collinear when $p=2$ and are non-collinear when p is an odd prime. Thus an important and simple geometric fact sharply distinguishes between the two named cases of these finite geometries.

This difference in the geometries (for the two cases) must be reflected in an important way in the theory of Abelian groups of order p^m and type $(1, 1, 1, \cdots)$. Early in the development of the theory of these groups it was noticed that their properties differ owing to whether p is 2 or is an odd prime. From our geometric interpretation and the facts just stated (in this section), the fundamental basis for this difference is apparent. Hence, in investigating these groups, one will now know precisely from what place to begin to develop those features of the theory which depend on the even or odd character of p.

For the case of the Abelian group $G_{(k+1)n}$, with the geometry $PG(k, p^n)$ constructed from it, the distinguishing difference of the two cases may be stated in group-theory language as follows (it being assumed now that $k>1$): Let A, B, C, D be four subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is contained in the group generated by another two while D is contained in the group $\{A, B, C\}$. Let E be the (unique) subgroup of the normal set common to the groups $\{A, B\}$ and $\{C, D\}$, F that common to the groups $\{A, C\}$ and $\{B, D\}$ and G that common to the groups $\{A, D\}$ and $\{B, C\}$. Then each of the subgroups E, F, G is in the subgroup generated by the other two when and only when $p=2$.

A large part of the theory of the geometry $PG(k, p^n)$ can be developed independently of any hypothesis as to the collinearity or noncollinearity of the diagonal points of a complete quadrangle (see § 6 of this paper). These theorems will give rise to corresponding theorems about Abelian groups of order p^m and type $(1, 1, 1, \cdots)$ which are independent of the odd or even character of p.

5. *The Theorems of Desargues and Pascal.* As an example of another interesting theorem in the theory of groups obtained from a geometric fact, let us consider the following.

The theorem of Desargues, which is valid in the $PG(k, p^n)$, may be stated thus. Let ABC and abc be two triangles in the same plane and let them be perspective from a point O so that O, A, a are collinear, O, B, b are collinear, and O, C, c are collinear. Let γ be the point of intersection of AB and ab, β that of AC and ac, and α that of BC and bc. Then the points α, β, γ are collinear.

Let us translate this result into a theorem concerning the Abelian group

$G_{(k+1)n}$ viewed as indicated in § 2 in the light afforded by the geometry $PG(k, p^n)$, it being assumed now that $k > 1$.

Let A, B, C be three subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is in the group generated by the other two. We select other subgroups of the normal set as follows, each of them to be in the group $\{A, B, C\}$: O is any such subgroup which is not contained in any one of the groups $\{A, B\}$, $\{B, C\}$, $\{C, A\}$; a, b, c are such subgroups different from O, A, B, C and contained respectively in the groups $\{O, A\}$, $\{O, B\}$, $\{O, C\}$. Let γ, α, β be the subgroups of the normal set of sub-subgroups common to the respective pairs of groups

$$\{A, B\},\ \{a, b\};\quad \{B, C\},\ \{b, c\};\quad \{C, A\},\ \{c, a\}.$$

Then each of the subgroups α, β, γ is in the subgroup generated by the other two.

The generalizations of the theorem of Desargues to higher dimensions yield likewise interesting theorems concerning Abelian groups. As phrased abstractly the theorems seem to be rather complicated; but in their geometric formulation they are easily comprehended and retained in mind.

As affording a final illustration of this method of translating geometric theorems into theorems about Abelian groups, let us consider the following which gives rise to the configuration of Pappus (Veblen and Young, *Projective Geometry*, Vol. I, p. 98). If A, B, C are any three distinct points of a line l, and A', B', C', are any three additional distinct points on another line l' meeting l in O, the three points γ, α, β of intersection of the respective pairs of lines

$$AB', A'B;\quad BC', B'C;\quad CA', C'A$$

are collinear.

Translating as in the previous case we have the following theorem:

Let O, A, A' be three subgroups of a normal set of subgroups of $G_{(k+1)n}$ such that no one of them is in the group generated by the other two. Let B and C be two additional subgroups contained in the group $\{O, A\}$ and belonging to the normal set, and B' and C' be two additional such subgroups contained in the group $\{O, A'\}$, these groups being existent when and only when $p^n > 2$ and $k > 1$. Let γ, α, β be the subgroups of the normal set which are common to the respective pairs of groups

$$\{A, B'\}, \{A', B\};\quad \{B, C'\}, \{B', C\};\quad \{C, A'\}, \{C', A\}.$$

Then each of the subgroups α, β, γ is in the subgroup generated by the other two.

6. *Geometries Affording Applications to Abelian Groups.* The analysis and development of projective geometry given by O. Veblen and J. W. Young (*Projective Geometry,* Vol. I, 1910; Vol. II, 1918) afford a convenient means of ascertaining what geometries have direct applications to the theory of Abelian groups by means of the representations of finite geometries given in the foregoing pages. In vol. II (p. 36) of this work, Veblen describes nine classes of geometries characterized by means of the assumptions which underlie them. Using capital letters to denote the assumptions and employing the notation of Veblen and Young (see the index to vol. II under the word "Assumption"), we select for our purpose four of these geometries as follows: A space satisfying Assumptions

A, E	is a general projective space;
A, E, P	is a proper projective space;
$A, E, \bar{H}$	is a modular projective space;
$A, E, \bar{H}, Q$	is a rational modular projective space.

It is easy to verify that the assumptions involved in these four geometries are all valid in the case of the geometry $PG(k, p^n)$, except that Q is valid when and only when $n = 1$. Since the points of this geometry have been represented by certain subgroups of the Abelian group $G_{(k+1)n}$, it follows that every theorem in any one of the four geometries named is capable of immediate translation into a theorem concerning the given Abelian group. In many cases a single theorem is capable of being so translated in a variety of ways, there being at least one such translation for every factorization of the number $(k+1)n$ into a product of two factors $k+1$ and n such that k and n are positive integers.

Each of the four geometries may be divided into two parts. In one part we have the assumption H_0, namely:

H_0. The diagonal points of a complete quadrangle are noncollinear. In the other we have the assumption that these diagonal points are collinear. The consequences of this latter assumption are not developed in detail by Veblen and Young, but many of the theorems given as dependent on A, E, P, H_0 (so far as the given proofs go) are provable without the use of H_0 (cf. vol. I, p. 261, exercise). We have seen (§ 4) that H_0 is valid in $PG(k, p^n)$ when and only when the prime p is different from 2.

Now in volume I of the work named no assumptions are used except those which are valid for $PG(k, p^n)$. Hence every theorem in volume I may be translated, in the way indicated, into a theorem about Abelian groups. The same remarks may be made about certain parts of volume II, and in

particular about chapter III and the first part of chapter IV. It is thus apparent that our representation of the $PG(k, p^n)$ by means of Abelian groups carries at once a large part of the results of projective geometry into the domain of Abelian groups and that they there become theorems about Abelian groups. Thus by a single act of thought a significant extension is given to the theory of Abelian groups and a method is made apparent by which the theory may be further developed.

Note on a Certain Generalization of the Preceding Results. Let us now consider more generally an Abelian group A whose order is a power of a prime p and whose type is $(m_1, m_2, \cdots, m_{(k+1)n})$. Let us denote a set of independent generators of A by

$$\begin{array}{l} a_{01}, a_{02}, a_{03}, \cdots, a_{0n}, \\ a_{11}, a_{12}, a_{13}, \cdots, a_{1n}, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{k1}, a_{k2}, a_{k3}, \cdots, a_{kn}, \end{array}$$

these being chosen so that a_{ij} is of period $p^{m_{in+j}}$. Then every element of A may be represented uniquely in the form

$$\prod_{i=0}^{k} a_{i1}^{s_{i1}} a_{i2}^{s_{i2}} \cdots a_{in}^{s_{in}}$$

where the exponent s_{ij} is a number of the set $0, 1, 2, \cdots, p^{m_{in+j}} - 1$.

Consider the following subset of these elements, namely,

$$\prod_{i=0}^{k} a_{i1}^{\sigma_{i1}} a_{i2}^{\sigma_{i2}} \cdots a_{in}^{\sigma_{in}}$$

where each σ runs over the set $0, 1, 2, \cdots, p-1$, or more generally the exponent σ_{ij} runs over the set $lp^{\alpha_{ij}}$ for $l = 0, 1, 2, \cdots, p-1$, the fixed integer α_{ij} being non-negative and less than m_{in+j}. An element of this sort, for the fixed set of exponents σ_{ij},

$$\sigma_{ij} = l_{ij} p^{\alpha_{ij}},$$

the a_{ij} and the α_{ij} having been chosen once for all, may be uniquely represented by the symbol

$$\{\mu_0, \mu_1, \cdots, \mu_k\}$$

where μ_i $(i = 0, 1, 2, \cdots, k)$ denotes that mark of the Galois field $GF[p^n]$ which may be written in the form

$$\mu_i = l_{i1} + l_{i2}\omega + l_{i3}\omega^2 + \cdots + l_{in}\omega^{n-1},$$

ω being a fixed primitive mark of $GF[p^n]$.

Now let $\mu_0, \mu_1, \cdots, \mu_k$ be a fixed set of $k+1$ marks of the field $GF[p^n]$, at least one of them being different from zero; and consider the set of elements

$$\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$$

where μ is a variable running over the p^n-1 non-zero marks of $GF[p^n]$. These elements generate a certain subgroup of A which we denote by the symbol $(\mu_0, \mu_1, \cdots, \mu_k)$. The same subgroup is denoted by the symbol $(\sigma\mu_0, \sigma\mu_1, \cdots, \sigma\mu_k)$ where σ is any non-zero mark of $GF[p^n]$. The total set of such subgroups we will call a normal set of subgroups of A.

The subgroups each of which is denoted by a symbol of the type $(\mu_0, \mu_1, \cdots, \mu_k)$ will be taken as the points of the geometry we are constructing. The point corresponding to the subgroup $(\mu_0, \mu_1, \cdots, \mu_k)$ will be denoted by the symbol $(\mu_0, \mu_1, \cdots, \mu_k)$, and $\mu_0, \mu_1, \cdots, \mu_k$ will be called the homogeneous coördinates of the point. In the geometry thus constructed the points are denoted by the same symbols as those employed in § 2 in constructing the geometry $PG(k, p^n)$ and the number system (the Galois field $GF[p^n]$) bears the same relation to the geometry in the new case as in the old. Hence the two geometries are abstractly the same. That is to say, the geometry constructed in this note is but another concrete representation of the abstract geometry $PG(k, p^n)$.

From this it follows that certain properties of the group A in the general case are identical with those for the special case when the type is $(1, 1, 1, \cdots)$, namely, those properties which may be expressed in terms of the points (and classes of points—lines, etc.) of the geometry $PG(k, p^n)$. For the sake of simplicity we shall deal with the special case when the group is of type $(1, 1, 1, \cdots)$; but the results will have the obvious extension indicated.

II. Groups of Isomorphisms of Abelian Groups of Type $(1, 1, 1, \cdots)$.

7. *Relation between the Groups $GLH\{k+1, p^n\}$ and I.* Let $G_{(k+1)n}$ as before be an Abelian group of order $p^{(k+1)n}$ and type $(1, 1, 1, \cdots)$. We denote it more simply by G when there is no danger of confusion. Let I denote the group of isomorphisms of G. As in the earlier part of § 2 we denote an element of this group by the symbol $\{x_0, x_1, x_2, \cdots, x_k\}$ where $x_0, x_1, \cdots, x_k$ are marks of the Galois field $GF[p^n]$.

Let us consider a linear homogeneous transformation

$$x_i' = \sum_{j=0}^{k} a_{ij}x_j, \qquad (i = 0, 1, 2, \cdots, k),$$

on the marks of this symbol, the coefficients a_{ij} being marks of $GF[p^n]$ and

the determinant $| a_{ij} |$ of this transformation being different from zero. If $\{x_0, x_1, \cdots, x_k\}$ runs over all the elements of the group G it is clear that $\{x_0', x_1', \cdots, x_k'\}$ likewise runs over all these elements. The transformation thus establishes a one-to-one correspondence of the elements of the group to its elements in some order. In each of these the identity corresponds to itself. Moreover, if $\{\mu_0, \mu_1, \cdots, \mu_k\}$ and $\{\nu_0, \nu_1, \cdots, \nu_k\}$ corresponds respectively to $\{\mu_0', \mu_1', \cdots, \mu_k'\}$ and $\{\nu_0', \nu_1', \cdots, \nu_k'\}$, then the product $\{\mu_0 + \nu_0, \cdots, \mu_k + \nu_k\}$ of the first pair of elements corresponds to the product $\{\mu_0' + \nu_0', \cdots, \mu_k' + \nu_k'\}$ of the corresponding (second) pair. Hence the correspondence of elements brought about by the given linear substitution effects an isomorphism of the group with itself. It is obvious that two distinct transformations effect different isomorphisms. Now the totality of linear homogeneous transformations of the given type constitutes the general linear homogeneous group $GLH\{k + 1, p^n\}$ on $k + 1$ indices with coefficients in the Galois field $GF[p^n]$. It is well known (and easily proved) that the order of this group is

$$(p^{(k+1)n} - 1)(p^{(k+1)n} - p^n)(p^{(k+1)n} - p^{2n}) \cdots (p^{(k+1)n} - p^{kn}).$$

This is a factor of the order

$$(p^{(k+1)n} - 1)(p^{(k+1)n} - p)(p^{(k+1)n} - p^2) \cdots (p^{(k+1)n} - p^{(k+1)n-1})$$

of the group I of isomorphisms of G; and it is a proper factor except when $n = 1$. Hence we have a proof of the known result that $GLH\{k + 1, p^n\}$ is a subgroup of I; it is a proper subgroup when and only when $n > 1$.

Let us consider more closely isomorphisms of G with itself which are effected by the named $GLH\{k + 1, p^n\}$. Let $\{\mu_0, \mu_1, \cdots, \mu_k\}$ be any element of G other than the identity and let $\{\mu_0', \mu_1', \cdots, \mu_k'\}$ be the element to which it corresponds under a given substitution belonging to $GLH\{k + 1, p^n\}$. Then the element $\{\mu\mu_0, \mu\mu_1, \cdots, \mu\mu_k\}$ corresponds to the element $\{\mu\mu_0', \mu\mu_1', \cdots, \mu\mu_k'\}$ under the same substitution. Hence the subgroup $(\mu_0, \mu_1, \cdots, \mu_k)$ corresponds to the subgroup $(\mu_0', \mu_1', \cdots, \mu_k')$. Therefore every substitution in the group $GLH\{k + 1, p^n\}$ effects an isomorphism of G with itself such that every subgroup of the corresponding normal set of subgroups corresponds to a subgroup of this set. Moreover, the multiplication of each coefficient a_{ij} in the transformation by one and the same non-zero mark ρ of the field gives a new transformation in which the correspondence of subgroups of the normal set as subgroups is unaltered while any other modification of the transformation, resulting in another transformation belonging to the group $GLH\{k + 1, p^n\}$ leads to a different correspondence of the subgroups as such.

Now the group $GLH\{k+1, p^n\}$ has $(p^n - 1, 1)$ isomorphism with the group $P(k, p^n)$ formed from the substitutions in $GLH\{k+1, p^n\}$ by treating $x_0, x_1, \cdots, x_k$ as the homogeneous coördinates in $PG(k, p^n)$, so that a substitution is now unchanged by multiplying each of its coefficients by one and the same non-zero mark ρ of the field. This group $P(k, p^n)$ is the projective group in $PG(k, p^n)$. From the result of the previous paragraph it follows that each substitution of the group $P(k, p^n)$ carries a subgroup of the normal set of subgroups into such a subgroup. Expressed geometrically this means that it transforms among themselves the points of the $PG(k, p^n)$.

When viewed geometrically, it is obvious that the group $P(k, p^n)$ also transforms planes into planes, 3-spaces into 3-spaces, and so on—facts which might be expressed also in the language of group theory. Thus a given substitution of $P(k, p^n)$ makes any given group generated by two subgroups of a normal set correspond to a group generated by two such subgroups; it also makes any given subgroups generated by three subgroups of the normal set correspond to a subgroup generated by three such subgroups; and so on.

8. *Analytical Representations of the Group I of Isomorphisms of G.* Let us consider the more general transformation

$$x_i' = \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} x_j^{p^{n-s}}, \qquad (i = 0, 1, 2, \cdots, k),$$

where the coefficients a_{ijs} are marks of $GF[p^n]$ such that these transformation equations have a unique solution for the symbols x_i in terms of the symbols x_i'. If

$$x_i' = \sum_{s=1}^{n} \sum_{j=0}^{k} b_{ijs} x_j^{p^{n-s}}, \qquad (i = 0, 1, 2, \cdots, k),$$

is a second transformation of the same kind, then the product of the two may be written in the form

$$\begin{aligned} x_i' &= \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} \Big(\sum_{\sigma=1}^{n} \sum_{\lambda=0}^{k} b_{j\lambda\sigma} x_\lambda^{p^{n-\sigma}}\Big)^{p^{n-s}} \\ &= \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} \Big(\sum_{\sigma=1}^{n} \sum_{\lambda=0}^{k} b_{j\lambda\sigma}^{p^{n-s}} x_\lambda^{p^{2n-s-\sigma}}\Big) \\ &= \sum_{\sigma=1}^{n} \sum_{\lambda=0}^{k} \sum_{s=1}^{n} \sum_{j=1}^{k} a_{ijs} b_{j\lambda\sigma}^{p^{n-s}} x_\lambda^{p^{2n-s-\sigma}} \\ &= \sum_{t=1}^{n} \sum_{l=0}^{k} \alpha_{ilt} x_l^{p^{n-t}}, \qquad (i = 0, 1, 2, \cdots, k), \end{aligned}$$

the α's being defined in a way which is obvious from a comparison of the last two members of the equation in the light of the fact that $x_\lambda^{p^n} = x_\lambda$. Thus

the product of two transformations of the class in consideration belongs also to the class. The named class of transformations therefore constitutes a group. This we shall call the group T. We shall prove that T, when interpreted as in the next paragraph, is identical with the group I of isomorphisms of G with itself. This result is known already for the case $k=0$ and for the case $n=1$.

Let $\{\mu_0', \mu_1', \cdots, \mu_k'\}$ and $\{\nu_0', \nu_1', \cdots, \nu_k'\}$ be the elements corresponding to $\{\mu_0, \mu_1, \cdots, \mu_k\}$ and $\{\nu_0, \nu_1, \cdots, \nu_k\}$ respectively under the given transformation with coefficients a_{ijs}. Then under the same transformation we have

$$\mu_i' + \nu_i' = \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs}(\mu_j^{p^{n-s}} + \nu_j^{p^{n-s}})$$

$$= \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs}(\mu_j + \nu_j)^{p^{n-s}}, \qquad (i=0, 1, 2, \cdots, k).$$

Hence $\{\mu_0' + \nu_0', \cdots, \mu_k' + \nu_k'\}$ corresponds to $\{\mu_0 + \nu_0, \cdots, \mu_k + \nu_k\}$ under the same transformation. Thence we see that if two given elements of G correspond respectively to two other given elements of G under a given transformation of T, then under the same transformation the product of the first pair of elements of G corresponds to the product of the second pair. Hence the substitution sets up an isomorphism of G with itself. Hence T is contained in the group I of isomorphisms of G. It remains to show that every element of I is in T.

For the latter purpose it is convenient to represent the group T in a different form.* Let ω be a primitive mark of $GF[p^n]$. Then any mark of $GF[p^n]$ may be written in the form

$$\gamma_0 + \gamma_1\omega + \gamma_2\omega^2 + \cdots + \gamma_{n-1}\omega^{n-1}$$

where each γ_i is a mark of $GF[p]$ and hence is an integer taken modulo p. Then we may write

$$x_i = \sum_{\lambda=0}^{n-1} \xi_{i\lambda}\omega^\lambda, \quad x_i' = \sum_{\lambda=0}^{n-1} \xi'_{i\lambda}\omega^\lambda, \quad a_{ijs} = \sum_{\lambda=0}^{n-1} a_{ijs\lambda}\omega^\lambda,$$

where the $\xi_{i\lambda}$, $\xi'_{i\lambda}$, $a_{ijs\lambda}$ are integers taken modulo p. Then the transformation τ of T, which has the coefficients a_{ijs}, may be written in the form

* The argument here is similar to that employed on pp. 69-70 of Dickson's *Linear Groups*.

$$\begin{aligned}\sum_{\lambda=0}^{n-1} \xi'_{i\lambda}\omega^{\lambda} &= \sum_{s=1}^{n}\sum_{j=0}^{k}\sum_{\lambda=0}^{n-1} a_{ijs\lambda}\omega^{\lambda}\Big(\sum_{\mu=0}^{n-1}\xi_{j\mu}\omega^{\mu}\Big)^{p^{n-s}} \\ &= \sum_{s=1}^{n}\sum_{j=0}^{k}\sum_{\lambda=0}^{n-1}\sum_{\mu=0}^{n-1} a_{ijs\lambda}\xi_{j\mu}{}^{p^{n-s}}\omega^{\mu p^{n-s}+\lambda} \\ &= \sum_{s=1}^{n}\sum_{j=0}^{k}\sum_{\lambda=0}^{n-1}\sum_{\mu=0}^{n-1} a_{ijs\lambda}\xi_{j\mu}\omega^{\mu p^{n-s}+\lambda}, \\ & \qquad\qquad (i=0,1,2,\cdots,k).\end{aligned}$$

Now every power of ω can be expressed linearly in terms of $\omega^0, \omega^1, \omega^2, \cdots, \omega^{n-1}$ with coefficients which are integers taken modulo p, since ω satisfies an equation of degree n with coefficients which are integers taken modulo p. On effecting this reduction we may write the last equation in the form

$$\sum_{s=1}^{n-1} \xi'_{i\sigma}\omega^{\sigma} = \sum_{\mu=0}^{n-1}\sum_{\sigma=1}^{n-1}\sum_{j=0}^{k} \alpha_{ij\mu\sigma}\xi_{j\mu}\omega^{\sigma}, \quad (i=0,1,2,\cdots,k),$$

where the $\alpha_{ij\mu\sigma}$ are integers taken modulo p. Equating coefficients of like powers of ω we have

$$\xi'_{i\lambda} = \sum_{\mu=0}^{n-1}\sum_{j=0}^{k} \alpha_{ij\mu\lambda}\xi_{j\mu}, \ (i=0,1,2,\cdots,k; \quad \lambda=0,1,2,\cdots,n-1).$$

Thus we have a linear transformation on the $(k+1)n$ quantities $\xi_{i\lambda}$, the coefficients of the transformation being integers taken modulo p. Since the x_i are uniquely expressible in terms of the x_i' it follows that the $\xi_{i\lambda}$ are uniquely expressible in terms of the $\xi'_{i\lambda}$ and thence that the transformation on the ξ's is non-singular.

Now the totality of such linear transformations on the $\xi_{i\lambda}$ is simply isomorphic with the group I of isomorphisms of G, as we see from the result at the end of the second paragraph of § 7 with n taken equal to 1. Hence in order to complete the proof that T is the group of isomorphisms of G it is sufficient to prove that each non-singular transformation on the $\xi_{i\lambda}$, such as the foregoing, is equivalent to a corresponding transformation in T.

In order to attain this end let the last foregoing transformation now be any non-singular linear transformation on the $\xi_{i\lambda}$ with coefficients which are integers taken modulo p. Change λ to σ, in the resulting equation (for fixed σ) multiply both sides by ω^{σ}, then sum as to σ from 0 to $n-1$. Thus we have the next preceding system of equations. From it we can go to the one which next precedes it provided that we are able to write

$$\sum_{\mu=0}^{n-1}\sum_{\sigma=1}^{n-1}\sum_{j=0}^{k} \alpha_{ij\mu\sigma}\xi_{j\mu}\omega^{\sigma} = \sum_{s=1}^{n}\sum_{j=0}^{k}\sum_{\lambda=0}^{n-1}\sum_{\mu=0}^{n-1} a_{ijs\lambda}\xi_{j\mu}\omega^{\mu p^{n-s}+\lambda}, \quad (i=0,1,2,\cdots,k),$$

where the coefficients $a_{ijs\lambda}$ are integers taken modulo p. If we have this

relation we can readily continue the reverse transformations through the equations written till we reach a transformation in the group T and having the coefficients a_{ijs}, these being marks in $GF[p^n]$. Hence, in order to show that every non-singular linear transformation on the $\xi_{i\lambda}$ (of the type in consideration) leads to a transformation of the group T it is sufficient to prove the existence of the integers $a_{ijs\lambda}$ modulo p such that the last foregoing system of equations reduces to an identity in the $\xi_{j\mu}$. For this purpose it is necessary and sufficient to show that integers $a_{ijs\lambda}$ modulo p exists such that the equation

$$\sum_{\sigma=0}^{n-1} \alpha_{ij\mu\sigma}\omega^\sigma = \sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} a_{ijs\lambda}\omega^{\mu p^{n-s}+\lambda}$$

is valid for each set of values i, j, μ. Let us write

$$\omega^{\mu p^{n-s}+\lambda} = \sum_{\sigma=0}^{n-1} \rho_{\mu\sigma s\lambda}\omega^\sigma,$$

where the coefficients $\rho_{\mu\sigma s\lambda}$ are integers taken modulo p. Then for the existence of the quantities $a_{ijs\lambda}$ it is necessary and sufficient that we have the relations

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} \rho_{\mu\sigma s\lambda}\, a_{ijs\lambda} = \alpha_{ij\mu\sigma}$$

for every i, j, μ, σ. If i and j are held fixed, these become n^2 equations in the n^2 unknown quantities $a_{ijs\lambda}$, $s = 1, 2, \cdots, n$, $\lambda = 0, 1, \cdots, n-1$. In order that they shall have a solution it is sufficient that their determinant D shall not vanish modulo p.

In order to prove that D does not vanish modulo p we shall show that we are led to a contradiction if we suppose that $D \equiv 0 \bmod p$. If $D \equiv 0 \bmod p$ then integers $t_{s\lambda}$ exist, not all congruent to zero modulo p, such that

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} t_{s\lambda}\rho_{\mu\sigma s\lambda} = 0, \qquad (\mu, \sigma = 0, 1, 2, \cdots, n-1).$$

For fixed σ multiply both members by ω^σ; then, summing as to σ we have a result which may be put in the form

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{n-1} t_{s\lambda} \sum_{\sigma=0}^{n-1} \rho_{\mu s\lambda}\omega^\sigma = 0, \qquad (\mu = 0, 1, 2, \cdots, \mu - 1);$$

or, in view of the definition of the quantities ρ,

$$\sum_{s=1}^{n} \sum_{\lambda=0}^{1-u} t_{s\lambda}\omega^{\mu p^{n-s}+\lambda} = 0;$$

or,

$$\sum_{s=1}^{n} \omega^{\mu p^{n-s}} \left(\sum_{\lambda=0}^{n-1} t_{s\lambda} \omega^{\lambda} \right) = 0, \qquad (\mu = 0, 1, \cdots, n-1).$$

Now no given one of the sums in the parenthesis can be zero unless every $t_{s\lambda}$ in that sum is zero. Hence, since not every $t_{s\lambda}$ is zero, one at least of these sums in the parenthesis is different from zero. Then the consistency of the foregoing system of equations requires that the determinant

$$\Delta = \begin{vmatrix} 1 & 1 & 1 & \cdots\cdots & 1 \\ \omega^{p^{n-1}} & \omega^{p^{n-2}} & \omega^{p^{n-3}} & \cdots\cdots & \omega^{p} \\ \omega^{2p^{n-1}} & \omega^{2p^{n-2}} & \omega^{2p^{n-3}} & \cdots\cdots & \omega^{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \omega^{(n-1)p^{n-1}} & \omega^{(n-1)p^{n-2}} & \omega^{(n-1)p^{n-3}} & \cdots\cdots & \omega^{(n-1)p} \end{vmatrix}.$$

shall vanish. But this determinant is, apart from a constant factor, equal to a product of factors each of which is of the form

$$\omega^{p^{n-\alpha}} - \omega^{p^{n-\beta}}, \qquad (\alpha, \beta = 1, 2, \cdots, n-1, \quad \alpha \neq \beta).$$

But, since ω is a primitive mark of $GF[p^n]$, no one of these factors can vanish. Hence $\Delta \neq 0$ in $GF[p^n]$. We have been led to this contradiction by assuming that $D \equiv 0 \bmod p$. Hence this congruence is not valid.

Summing up the argument, we have the following result:

The group T, defined and interpreted at the beginning of this section, is identical with the group I of isomorphisms of the Abelian group G of type $(1, 1, 1, \cdots)$.

If the group G is of order p^m then we have a different analytical representation of the group I of isomorphisms of G for each factorization of m in the form $(k+1)n$. For the group I itself we have the simplest representation when $n = 1$. The different possible representations, however, will furnish varying information (as we shall see later) concerning various subgroups of I.

9. *On Certain Subgroups of I.* When the group I of isomorphisms of G is written in the form of the transformation group T in $GF[p^n]$, certain interesting classes of subgroups become obvious. To construct the first one of these classes we proceed as follows. Let d be any divisor of n and write $n = d\nu$. Then in the typical transformation of T put a_{ijs} equal to zero when s is not divisible by d. Then the transformation takes the special form

$$x_i' = \sum_{t=1}^{\nu} \sum_{j=0}^{k} \bar{a}_{ijt} x_j^{p^{d(\nu-t)}}, \qquad (i = 0, 1, 2, \cdots, k).$$

The product of this transformation by another of the same form may be written as a transformation of this form, the method of reduction being the same as that employed at the beginning of § 8. The named transformations therefore form a group T_d which is a subgroup of the group T, and hence (under the interpretation used in § 8) a subgroup of the group I of isomorphisms of G. It is obvious that T_1 is identical with T.

The group T_d thus formed is a generalization of the Betti-Mathieu group (see Dickson's *Linear Groups*, pp. 64-70). Just as the Betti-Mathieu group may be identified with a linear homogeneous group (Dickson, *l. c.*, p. 69), so can its generalization T_d be similarly identified with a like group. The argument is a generalization of one employed in the preceding section, whence it is sufficient merely to outline it.

Let ω be a primitive mark of $GF[p^n]$. Then any mark of $GF[p^n]$ may be written in the form

$$\gamma_0+\gamma_1\omega+\cdots+\gamma_{\nu-1}\omega^{\nu-1}$$

where each γ_i is a mark of $GF[p^d]$. Then we may write

$$x_i=\sum_{\lambda=0}^{\nu-1}\xi_{i\lambda}\omega^\lambda,\quad x_i'=\sum_{\lambda=0}^{\nu-1}\xi'_{i\lambda}\omega^\lambda,\quad \bar{a}_{ijt}=\sum_{\lambda=0}^{\nu-1}\bar{a}_{ijt\lambda}\omega^\lambda,$$

where the $\xi_{i\lambda}$, $\xi'_{i\lambda}$, $\bar{a}_{ijt\lambda}$ are marks of $GF[p^d]$. The argument now proceeds in the same way as in the previous case and we find that

$$\xi'_{i\lambda}=\sum_{\mu=0}^{\nu-1}\sum_{j=0}^{k}\bar{\alpha}_{ij\mu\lambda}\xi_{j\mu},\qquad (i=0,1,\cdots,k,\quad \lambda=0,1,\cdots,\nu-1),$$

where the $\bar{\alpha}_{ij\mu\lambda}$ are marks of $GF[p^d]$. Thus a given transformation in T_d can be put into the form just written. Conversely, any transformation of the latter form can be put into the form of a transformation of T_d, the method of proof being that employed in the preceding section.

We have thus exhibited the group T_d as a homogeneous linear group in the Galois field $GF[p^d]$.

We shall now determine certain subgroups of I yielding point transformations in $PG(k,p^n)$. Let us consider the transformation

$$x_i'=\sum_{j=0}^{k}\alpha_{ijt}x_j^{p^t},\qquad (i=0,1,\cdots,k),$$

belonging to the group T of § 8. On combining this transformation with the similar transformation

$$x_i'=\sum_{j=0}^{k}\beta_{ij\tau}x_j^{p^\tau},\qquad (i=0,1,\cdots,k),$$

we have

$$\begin{aligned} x_i' &= \sum_{j=0}^{k} \alpha_{ijt} \big(\sum_{\lambda=0}^{k} \beta_{j\lambda\tau} x_\lambda^{p^\tau} \big)^{p^t} \\ &= \sum_{\lambda=0}^{k} \sum_{j=0}^{k} \alpha_{ijt} \beta^{p^t}{}_{j\lambda\tau} x_\lambda^{p^{t+\tau}} \\ &= \sum_{\lambda=0}^{k} \gamma_{i\lambda} x_\lambda^{p^{t+\tau}}, \qquad (i=0,1,\cdots,k), \end{aligned}$$

where the exponent $\tau + t$, when not less than n, is to reduced modulo n.

From this it follows that the foregoing set of transformations forms a group Γ if the coefficients α_{ijt} are marks of $GF[p^n]$ and t varies over the set $0, 1, 2, \cdots, n-1$. If d is any divisor of n and t ranges over the multiples of d in the set $0, 1, 2, \cdots, n-1$ we have a subgroup Γ_d of the group Γ. Thus we have a group Γ_d for each divisor d of n. Evidently Γ_1 is the same as Γ. We denote by Γ_0 the group in which t has the value 0 alone, this being a linear group.

Now in any particular transformation of Γ the quantities x enter homogeneously. Hence Γ has $(p^n-1, 1)$ isomorphism with the group of point transformations which it generates in $PG(k, p^n)$. This group of point-transformations we shall denote by $C(k, p^n)$. The subgroup corresponding to the subgroup Γ_d of Γ we shall denote by $C_d(k, p^n)$. The groups $C_d(k, p^n)$ are groups of point-transformations in $PG(k, p^n)$. The group $C_0(k, p^n)$ is identical with the projective group $P(k, p^n)$ which we encountered in § 7. We shall return in § 12 to a further study of these groups.

10. *The Holomorph of G.* The set of transformations of the form

$$x_i' = x_i + a_i, \qquad (i = 0, 1, \cdots, k),$$

where the a_i are marks of $GF[p^n]$, clearly form an Abelian group $\bar{G}$ of order $p^{(k+1)n}$ and type $(1, 1, 1, \cdots)$. It is therefore simply isomorphic with the given Abelian group G and may be taken as a representation of it. The group generated by this group and group T of § 8 is therefore a representation of the holomorph of G—a fact which generalizes a well-known result (see for instance Burnside's *Theory of Groups,* 2nd ed'n, p. 245). The holomorph of G may therefore be represented by the set of non-singular transformations each of which has the form

$$x_i' = \sum_{s=1}^{n} \sum_{j=0}^{k} a_{ijs} x_j^{p^{n-s}} + a_i, \qquad (i = 0, 1, \cdots, k),$$

where the a's are marks of $GF[p^n]$. For $n = 1$ this is a well-known result. The transformation group so defined will be represented by the symbol H.

It is well-known that the group $\bar{G}$ is a self-conjugate subgroup of H. It is therefore a self-conjugate subgroup of every subgroup of H which contains $\bar{G}$. In particular $\bar{G}$ is transformed into itself by the group T_d defined in § 9. Hence the group $\{T_d, \bar{G}\}$ is a subgroup of H of the same index as that T_d in T. We thus have a ready means of constructing it. Certain of its subgroups are obvious, namely those of the form $\{T_d, \bar{G}_i\}$ where $\bar{G}_i$ is a subgroup of $\bar{G}$. An analytical representation of $\{T_d, \bar{G}\}$ is afforded by the set of non-singular transformations

$$x_i' = \sum_{t=1}^{\nu} \sum_{j=0}^{k} \bar{a}_{ijt} x_j^{p^{d(\nu-t)}} + a_i, \qquad (i = 0, 1, \cdots, k),$$

where the a's are marks of $GF[p^n]$ and d is any factor of n.

Again we can form other subgroups of H in a similar manner by taking the groups Γ_d of § 9 and combining each of them with $\bar{G}$. The forms of the analytical representations of these groups are obvious.

11. *Certain Homogeneous Groups Suggested by T and H.* At the end of § 2 we saw that the points of the Euclidean finite geometry $EG(k+1, p^n)$ may be identified with the elements of the group $G_{(k+1)n}$. Hence the group I of isomorphisms of $G_{(k+1)n}$ may be considered as a group of point transformations in $EG(k+1, p^n)$. This suggests the derivation of homogeneous groups from T and H and their subgroups and the interpretation of these in $PG(k+1, p^n)$. Accordingly we shall consider the homogeneous group whose transformations are of the form

$$x_i' = \sum_{j=0}^{k} a_{ij1} x_j^{p^{d(\nu-1)}} + \sum_{s=2}^{\nu} \sum_{j=0}^{k} a_{ijs} x_j^{p^{d(\nu-t)}} x_{k+1}^{p^{d(\nu-1)} - p^{d(\nu-t)}} + a_i x_k^{p^{d(\nu-1)}}, \qquad (i = 0, 1, 2, \cdots, k),$$

$$x'_{k+1} = x_{k+1}^{p^{d(\nu-1)}},$$

where d is any positive integral divisor of n and $n = d\nu$ (it being understood that the second summation in the first of these equations is to be omitted when $\nu = 1$). When one of the variables x_{k+1} and x'_{k+1} is 1 (or 0) the other has the same value. Therefore the given transformation transforms the points $(x_0, x_1, \cdots, x_k, 1)$ of $EG(k+1, p^n)$ according to the same permutation as that by which the corresponding transformation in $\{T_d, \bar{G}\}$ (obtained by replacing x_{k+1} by 1) transforms the elements of $G_{(k+1)n}$ when denoted by coördinates as in § 2. The fixed $PG(k, p^n)$, namely $x_{k+1} = 0$, is transformed by the foregoing substitution according to the substitution

$$x_i' = \sum_{j=0}^{k} a_{ij1} x_j^{p^{d(\nu-1)}}, \qquad (i = 0, 1, 2, \cdots, k),$$

it being assumed that the a_{ij1} are now such that this transformation is non-

singular. The total homogeneous group whose transformations are of the form of the first foregoing substitution on $x_0, x_1, \cdots, x_{k+1}$ we shall denote by $\bar{H}_d$; the subgroup in the transformations of which each a_i is zero we shall denote by $\bar{T}_d$. These are the transformation groups on the points of $PG(k+1, p^n)$ which are suggested by the non-homogeneous groups of §§ 8-10. The case when $d=1$ deserves special attention on account of its connection with the group of isomorphisms and the holomorph of $G_{(k+1)n}$.

It is obvious that there exists in $PG(k+1, p^n)$ a similar group for every k-space and corresponding Euclidean space in this $(k+1)$-space, such k-space in the new group playing the role which the k-space $x_{k+1}=0$ plays in the group as originally defined.

III. COLLINEATION GROUPS.

12. *The Collineation Group in $PG(k, p^n)$.* We shall now prove the following theorem concerning the collineation group in $PG(k, p^n)$ and certain of its subgroups.

THEOREM I. *The collineation group $C(k, p^n)$ in $PG(k, p^n)$ is represented analytically by the homogeneous transformations*

$$\text{(A)} \qquad \rho x_i' = \sum_{j=0}^{k} \beta_{ij\tau} x_j^{p^\tau}, \qquad (i=0, 1, \cdots, k, \quad \tau = 0, 1, \cdots, n-1),$$

where the $\beta_{ij\tau}$ are marks of $GF[p^n]$ such that the determinant

$$\Delta_\tau \equiv \begin{vmatrix} \beta_{00\tau} & \beta_{01\tau} & \cdots & \beta_{0k\tau} \\ \beta_{10\tau} & \beta_{11\tau} & \cdots & \beta_{1k\tau} \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{k0\tau} & \beta_{k1\tau} & \cdots & \beta_{kk\tau} \end{vmatrix}$$

is different from zero for each value of τ. Its order is n times the order of its projective subgroup $P(k, p^n)$, or $C_0(k, p^n)$, made up of those transformations of (A) in each of which $\tau=0$ and is therefore *

$$n(p^{(k+1)n}-1)p^{kn}(p^{kn}-1)p^{(k-1)n}(p^{(k-1)n}-1)\cdots p^{2n}(p^{2n}-1)p^n$$
$$=\{n/(p^n-1)\}\prod_{i=0}^{k}(p^{(k+1)n}-p^{in}).$$

The group C is generated by C_0 and the collineation

$$\rho x_i' = x_i^p, \qquad (i=0, 1, \cdots, k).$$

The last element transforms C_0 into itself.

---

* Compare Dickson's *Linear Groups*, p. 87. Our group $P(k, pn)$ is equivalent to the group of linear fractional transformations here treated by Dickson.

If d is any proper divisor of n, then we have a subgroup $C_d(k, p^n)$ of $C(k, p^n)$ (with $C_1 \equiv C$) generated by C_0 and the collineation

$$\rho x_i' = x_i^{p^d}, \qquad (i = 0, 1, \cdots, k);$$

and C_d is of index d in C. The transformations in C_d are of the form of (A) *with the restriction on τ that it shall be confined to the multiples of d belonging to the sequence $0, 1, \cdots, n-1$.*

Those transformations in C_d whose determinants Δ_τ are $(k+1)$-th *powers in $GF[p^n]$ form a subgroup $\bar{C}_d(k, p^n)$ of C_d of index μ where μ is the greatest common divisor of $k+1$* and *$p^n - 1$.*

The projective group $P(k, p^n)$, considered as a permutation group on the points of $PG(k, p^n)$, is triply transitive when $k = 1$ and is doubly transitive when $k > 1$. The same property of transitivity belongs to each of the previously named groups which contains $P(k, p^n)$ as a subgroup.

The group $\bar{C}_d(k, p^n)$ is doubly transitive, when considered as a permutation group on the points of $PG(k, p^n)$.

Finally, in a special case, we have another subgroup of C defined as follows. Let $k+1$ be a divisor of n, and let σ be a fixed divisor of $n/(k+1)$. Moreover, let $k+1$ be a factor of $p^\sigma - 1$. Any multiple of σ in the set $0, 1, \cdots, n-1$ can be written in just one way in the form $\{(k+1)s + \alpha\}\sigma$ where $0 \leqq \alpha \leqq k$ and s is a non-negative integer. For every such multiple of σ form the entire set of homogeneous transformations

$$\text{(B)} \qquad \rho x_i' = \sum_{j=0}^{k} \beta_{ijs\alpha} x_j^{p^{\{(k+1)s+\alpha\}\sigma}}, \qquad (i = 0, 1, \cdots, k),$$

in which each determinant $|\beta_{ijs\alpha}|$ of a transformation (s and α being fixed for a particular determinant) is equal to ω^α times a $(k+1)$-th *power in $GF[p^n]$, ω being a primitive mark of $GF[p^n]$. The totality of these transformations forms a subgroup $H_\sigma(k, p^n)$ of C which is of index $(k+1)\sigma$ in C. Moreover H_σ is contained in C_σ and is of index $k+1$ in C_σ. The group H_σ is generated by $\bar{C}_0$ and the transformations of the form*

$$\text{(C)} \qquad \rho x_i' = \omega^{t_i} x_i^{p^{\{(k+1)s+\alpha\}\sigma}}, \qquad (i = 0, 1, \cdots, k),$$

where $t_0 + t_1 + \cdots + t_k \equiv \alpha$ mod *$(k+1)$. When considered as a permutation group on the points of $PG(k, p^n)$, the group $H_\sigma(k, p^n)$ is triply transitive when $k = 1$ and is doubly transitive when $k > 1$.*

The collineation group described in the first paragraph of the theorem is the group to which we were led in § 9 in treating the subgroups of the group T of isomorphisms of G. It is an easy step to prove that the group is

a collineation group. To prove that it contains all collineations in $PG(k, p^n)$ is more difficult. But this has been effected by Veblen * through the aid of earlier work by Veblen and Bussey and by Levi. The result stated in the first paragraph of the theorem is therefore already known.

The proof of the statement in the second paragraph is omitted since it is almost immediate.

If two substitutions in C_d have their determinants equal to $(k+1)$-th powers, then their product has its determinant equal to a $(k+1)$-th power, as one may prove easily by combining these substitutions and making use of the fact that the p-th power of a determinant D whose elements are in $GF[p^n]$ is equal to the determinant $\bar{D}$ whose elements are the p-th powers of the corresponding elements of D. This proves the existence of the subgroup named in the third paragraph of the theorem. That this subgroup is of index μ in C_d is proved in general by the same method as that employed by Dickson (*l. c.*, p. 87) for the case of the group C_0.

The transitivity properties named in the fourth paragraph are immediate consequences of the fact that there exists in $P(k, p^n)$ a transformation which carries any $k+2$ points of $PG(k, p^n)$, no $k+1$ of which are on the same $(k-1)$-space, into any like set of $k+2$ points.

To show that $\bar{C}_d$ is doubly transitive we note first that the transformation (A) carries the points $(1, 0, 0, \cdots, 0)$ and $(0, 1, 0, \cdots, 0)$ into the points

$$(\beta_{00\tau}, \beta_{10\tau}, \cdots, \beta_{k0\tau}) \text{ and } (\beta_{01\tau}, \beta_{11\tau}, \cdots, \beta_{k1\tau})$$

respectively. Call these the points C and D respectively. The transformation may be chosen so that C and D are any two assigned points of $PG(k, p^n)$. Since C and D are different points there exist integers λ and μ such that the determinant $\beta_{\lambda 0\tau}\beta_{\mu 1\tau} - \beta_{\lambda 1\tau}\beta_{\mu 0\tau}$ is different from zero. Suppose now that $k > 1$. From the transformations (A) which carry the first named points into C and D respectively choose one as follows: take $\beta_{\lambda s\tau} = 0 = \beta_{\mu s\tau}$ for $s = 2, 3, \cdots, k$; choose the remaining $\beta_{ij\tau}$ for which $j > 1$ so as to give to the determinant Δ_τ any preassigned value different from zero. It is obvious that this can be done. Hence the choice of the β's and τ can be made so that the transformation (A) thus constructed belongs to the group $\bar{C}_d$. Hence the group $\bar{C}_d(k, p^n)$ is doubly transitive when $k > 1$. It is well known (cf. Dickson, *l. c.*, p. 261) and is easily proved that it is doubly transitive when $k = 1$. Hence the group $\bar{C}_d$ is doubly transitive in all cases.

It remains to prove the statements in the last paragraph of the theorem.

To show that the system of transformations named constitute a group,

* *Transactions of the American Mathematical Society*, Vol. 8 (1907), pp. 366-368.

consider two transformations of the named form, in one of which s and α are replaced by s_1 and α_1 and in the other of which they are replaced by s_2 and α_2. The product of these two transformations (in one order) may be written in the form

$$\rho x_i' = \sum_{j=0}^{k} \beta_{ijs_1\alpha_1} \left\{ \sum_{\mu=0}^{k} \beta_{j\mu s_2\alpha_2} x_\mu^{p^{\{(k+1)s_2+\alpha_2\}\sigma}} \right\}^{p^{\{(k+1)s_1+\alpha_1\}\sigma}}$$
$$= \sum_{\mu=0}^{k} \left\{ \sum_{j=0}^{k} \beta_{ijs_1\alpha_1} (\beta_{j\mu s_2\alpha_2})^{p^{\{(k+1)s_1+\alpha_1\}\sigma}} \right\} x_\mu^{p^{\{(k+1)(s_1+s_2)+\alpha_1+\alpha_2\}\sigma}}$$

for $i = 0, 1, \cdots, k$. It is easy to see that the determinant of this product transformation can be written as a product of determinants in the form

$$| \beta_{ijs_1\alpha_1} | \cdot | \beta_{j\mu s_2\alpha_2} |^{p^{\{(k+1)s_1+\alpha_1\}\sigma}}.$$

Now the exponent on the second determinant is congruent to 1 modulo $k+1$ since $p^\sigma - 1$ is divisible by $k+1$. Hence the determinant of the last written transformation is of the form of a $(k+1)$-th power in $GF[p^n]$ times $| \beta_{ijs_1\alpha_1} | \cdot | \beta_{j\mu s_2\alpha_2} |$. But these two determinants (by hypothesis) are equal to $(k+1)$-th powers in the $GF[p^n]$ times ω^{α_1} and ω^{α_2} respectively. Hence the determinant of the product transformation is equal to a $(k+1)$-th power in $GF[p^n]$ times $\omega^{\alpha_1+\alpha_2}$. From this and the fact that n is a multiple of $(k+1)\sigma$ it follows that the product transformation belongs to the set of transformations defined in the last paragraph of the theorem. That set therefore forms a group H_σ. It is obviously contained in C.

It is obvious that a general transformation (B) of H_σ may be multiplied by a suitable transformation (C) so as to produce a transformation belonging to $\bar{C}_0$. From this and the fact that every transformation (C) is in H_σ it follows readily that H_σ has the named generators.

It is obvious that $H_\sigma(k, p^n)$ is a subgroup of $C_\sigma(k, p^n)$. Moreover the general transformation in C_σ has its determinant restricted to be different from zero while a like transformation in H_σ has a further restriction that the value of its determinant shall be of a certain form relative to $(k+1)$-th powers so that the possible values for the determinants of transformations in C_σ of given form are $k+1$ times as many in number as the possible values for the determinants of the corresponding transformations in H_σ. From this it follows without difficulty that H_σ is of index $k+1$ in C_σ. It is therefore of index $(k+1)\sigma$ in C.

It remains to establish the transitivity properties of the group $H_\sigma(k, p^n)$. For the case $k=1$ the group was investigated by E. Mathieu.* In particular,

* *Journal de Mathématiques,* Ser. 2, Vol. 6 (1861), pp. 241-323.

he proved (p. 264) that it is triply transitive. Hence it remains to consider the case in which $k > 1$. In this case the same argument can be used as that by means of which the double transitivity of $\bar{C}_d$ was established and with the conclusion that $H_\sigma(k, p^n)$ is doubly transitive when $k > 1$.

This completes the proof of the theorem.

The transformation groups appearing in the foregoing theorem have been interpreted in it as permutation groups on the points of $PG(k, p^n)$. But these groups transform lines into lines; hence they transform the m-spaces $PG(m, p^n)$ contained in $PG(k, p^n)$ among themselves for each value m of the set $0, 1, 2, \cdots, k-1$. (Here we are taking k to be greater than 1.) Hence they may be interpreted as permutation groups on the symbols denoting the m-spaces for each particular value of m.

In particular, the $(k-1)$-spaces are transformed among themselves. The corresponding permutation group is of the same degree as that on the points of $PG(k, p^n)$, since the number of $(k-1)$-spaces in $PG(k, p^n)$ is equal to the number of points in this k-space. In view of the principle of duality it is not difficult to show that the two permutation groups arising from $C(k, p^n)$ are identical as permutation groups; for every transformation (A) on the points of $PG(k, p^n)$ can be expressed in the form of a transformation of the same general type on the coördinates which represent in a dual way the $(k-1)$-spaces $PG(k-1, p^n)$ in $PG(k, p^n)$. Moreover, the transformations (A) themselves set up a one-to-one correspondence among the elements of $C(k, p^n)$ when interpreted on the one hand as permutations on the points of $PG(k, p^n)$ and on the other hand as permutations on the $(k-1)$-spaces in $PG(k, p^n)$. Furthermore it may be seen that this correspondence is not the identical correspondence; for there are transformations leaving fixed the $(k-1)$-space $x_k = 0$ without leaving fixed any point of $PG(k, p^n)$. Detailed evidence of this fact will appear in the next section; it is involved in the fact that both the subspace $x_k = 0$ and the corresponding Euclidean space $EG(k, p^n)$ may have all its points permuted among themselves by one and the same transformation of $C(k, p^n)$.

The results of the last paragraph may be generalized to the case of l-spaces and their duals the $(k-l-1)$-spaces. Each of these sets of spaces is permuted by the transformations of $C(k, p^n)$ and the two permutation groups thus arising are identical as permutation groups. Again the simple isomorphism which is established between them is not the identical isomorphism, except in the special case of a self-dual set of spaces. This may be seen by observing that a space of the one type may be held fixed while no space of the other type is held fixed.

Hence we have the following theorem:

THEOREM II. *The collineation group $C(k, p^n)$ (when $k > 1$) transforms the $(k-1)$-spaces $PG(k-1, p^n)$ in $PG(k, p^n)$ according to the same permutation group as that according to which it transforms the points of $PG(k, p^n)$; it sets up a simple isomorphism of this permutation group with itself which is different from the identical isomorphism. More generally it sets up a like correspondence between two identical permutation groups the letters of one of which are the symbols for the l-spaces of $PG(k, p^n)$ while the letters of the other are the symbols for the dual $(k-l-1)$-spaces (except that the isomorphism may be identical in the case of self-dual spaces). These several permutation groups (of different degrees) are all simply isomorphic since each of them is simply isomorphic with $C(k, p^n)$ itself.*

It is obvious that similar results may be established for each of the subgroups of $C(k, p^n)$ described in theorem I. Of particular interest is the corresponding theorem for the case of the projective group $P(k, p^n)$. Thus *theorem* II *becomes a new theorem of interest if throughout it we replace $C(k, p^n)$ by $P(k, p^n)$ wherever the former occurs.*

For the case when $k > 1$ the lines of $PG(k, p^n)$ are permuted among themselves by $P(k, p^n)$, or $C(k, p^n)$, according to a transitive group, since any $k+2$ points no $k+1$ of which are on a $(k-1)$-space may be transformed into such a set of $k+2$ points by either of the named groups. If $k > 2$ the $PG(k, p^n)$ has pairs of intersecting lines and pairs of lines which do not intersect: since a pair of one of these sorts can not be transformed into a pair of the other sort, it follows that this permutation group on the lines of $PG(k, p^n)$ can not be doubly transitive when $k > 2$. When $k = 2$ the lines are transformed according to the same permutation group as the points, the latter being the dual of the former in this case. Hence the lines of $PG(2, p^n)$ are transformed among themselves according to a doubly transigroup both by $C(2, p^n)$ and by $P(2, p^n)$.

More generally it may be shown in the same way that the m-spaces $PG(m, p^n)$ in $PG(k, p^n)$, when $m < k$ and $k > 1$, are permuted according to a transitive group by either $P(k, p^n)$ or $C(k, p^n)$. If $0 < m < \frac{1}{2}k$ this group is simply transitive since there exist two sorts of pairs of m-spaces, namely, pairs in which the two spaces intersect and those in which they do not intersect, and a pair of one sort can not be transformed into a pair of the other sort by either group in consideration. Thence by means of the principle of duality it is seen that this permutation group is also simply transitive when $\frac{1}{2}k < m < k-1$. We have to consider further the case when k is even and

$m = \frac{1}{2}k$. Since this case has already been treated when $k = 2$, we shall now suppose that $k > 2$. Then for this case we have $m \geqq 2$. It is clear, then, that there exist again two sorts of pairs of m-spaces, namely, pairs in which the elements have an $(m-1)$-space in common and pairs in which the common elements constitute a space of fewer dimensions. Since a pair of one of these sorts can not be transformed into a pair of the other sort by either $P(k, p^n)$ or $C(k, p^n)$ we conclude in this case also that the permutation group on the m-spaces as symbols is simply transitive.

We shall now show that the permutation group generated in the m-spaces by $P(k, p^n)$, and hence that generated by $C(k, p^n)$, is primitive. Since the group is doubly transitive when $m = 0$ or $k - 1$ we may confine ourselves to the case in which $0 < m < k - 1$. We assume that the group is imprimitive and show that we are thus led to a contradiction. Since the m-spaces in any given $(m+1)$-space of $PG(k, p^n)$ are permuted among themselves in a doubly transitive way by the subgroup which leaves this $(m+1)$-space invariant, it follows that the m-spaces in any given $(m+1)$-space must all belong to the same set of imprimitivity. Thence it follows that the set of imprimitivity to which any given m-space M belongs must contain all the m-spaces included in the totality of $(m+1)$-spaces each of which contains M. If $m + 1 < k$ fix attention on all the $(m+1)$-spaces containing M and lying in one and the same $(m+2)$-space, and also all the $(m+1)$-spaces in this $(m+2)$-space and containing any m-space already obtained by this process of construction. Since every two $(m+1)$-spaces in the $(m+2)$-space contain an m-space in common it follows that the named process brings into consideration all the $(m+1)$-spaces in the given $(m+2)$-space. Hence every m-space in the $(m+2)$-space belongs to the same set of imprimitivity as M itself. If $m + 2 < k$ one can prove in a similar manner that the set of imprimitivity containing M contains also all m-spaces in a given $(m+3)$-space containing the given $(m+2)$-space; and so on. Hence the given set of imprimitivity contains all the m-spaces in $PG(k, p^n)$. Since this is impossible for a set of imprimitivity, we conclude that the permutation group in question is primitive.

Gathering up the results, we have the following theorem:

THEOREM III. *When $k > 1$ the collineation group $C(k, p^n)$, or its projective subgroup $P(k, p^n)$, transforms the m-spaces of $PG(k, p^n)$, $m < k$, according to a primitive permutation group; this group is doubly transitive when $m = 0$ or $k - 1$, otherwise it is simply transitive.*

From theorems II and III and from the groups $C_d(k, p^n)$ of theorem I we have the following theorem as an obvious corollary:

THEOREM IV. *There is no upper limit K to the number of primitive groups (of varying degrees) in a set of primitive groups each group of which is simply isomorphic with each of the others in the set. For every integer L there exist integers $s[t]$ such that the number of doubly transitive [triply transitive] groups of degree $s[t]$ is greater than L.*

13. *Collineation Groups Leaving Invariant an $EG(k, p^n)$.* The groups described in theorem I of § 12 obviously have corresponding subgroups each of which leaves invariant a $PG(k-1, p^n)$ in $PG(k, p^n)$. The points of $PG(k, p^n)$, not in a particular $PG(k-1, p^n)$ contained in it, form a Euclidean finite geometry of p^{kn} points; it is denoted by $EG(k, p^n)$. The named subgroups, leaving invariant a $PG(k-1, p^n)$, obviously transform among themselves the points of the corresponding $EG(k, p^n)$. Without real loss of generality we take the fixed $PG(k-1, p^n)$ to be that defined by the equation $x_0 = 0$. We then use $EG(k, p^n)$ for the corresponding Euclidean finite geometry. Concerning the named subgroups to which we are thus led, we have the following theorem which we shall now prove:

THEOREM I. *The collineation group $C(k, p^n)$ has a subgroup $EC(k, p^n)$ whose transformations may be represented analytically in the form*

$$(\bar{A}) \qquad \begin{aligned} \rho x_0' &= \beta_\tau x_0^{p^\tau}, & (\beta_\tau \neq 0),\\ \rho x_i' &= \sum_{j=0}^{k} \beta_{ij\tau} x_j^{p^\tau}, & (i = 1, 2, \cdots, k), \end{aligned}$$

where τ runs over the sequence $0, 1, 2, \cdots, n-1$. Its order is n times the order of its subgroup $EP(k, p^n)$, or $EC_0(k, p^n)$, made up of those transformations of $(\bar{A})$ in each of which $\tau = 0$ and is therefore

$$np^{kn} \prod_{i=0}^{k-1} (p^{kn} - p^{in}).$$

The group EC is generated by EC_0 and the collineation

$$\rho x_i' = x_i^p, \qquad (i = 0, 1, 2, \cdots, k).$$

The last element transforms EC_0 into itself.

If d is any proper divisor of n, then we have a subgroup $EC_d(k, p^n)$ of $EC(k, p^n)$ (with $EC_1 \equiv EC$) generated by EC_0 and the collineation

$$\rho x_i' = x_i^{p^d}, \qquad (i = 0, 1, 2, \cdots, k);$$

and EC_d is of index d in EC. The transformations in EC_d are of the form

of ($\bar{\text{A}}$) *with the restriction on* τ *that it shall be confined to the multiples of* d *belonging to the sequence* $0, 1, 2, \cdots, n-1$.

Those transformations in EC_d *whose determinants are* $(k+1)$*-th powers in* $GF[p^n]$ *form a subgroup* $E\bar{C}_d(k, p^n)$ *of* EC_d *of index* μ *where* μ *is the greatest common divisor of* $k+1$ *and* p^n-1.

The group $EP(k, p^n)$, *considered as a permutation group on the* p^{kn} *points of* $EG(k, p^n)$, *is doubly transitive. Moreover, it is triply transitive when* $k > 1$ *and* $p^n = 2$. *The same property of transitivity belongs to each of the previously named groups which contains* $EP(k, p^n)$ *as a subgroup.*

Considered as a permutation group on the p^{kn} *points of* $EG(k, p^n)$, *the group* $E\bar{C}_d(k, p^n)$ *is doubly transitive when* $k > 1$ *and also when* $k = 1$ *and* $p = 2$; *it is singly transitive when* $k = 1$ *and* p *is an odd prime. This singly transitive group is primitive.*

Finally, in a special case, we have another subgroup of EC *defined as follows. Let* $k+1$ *be a divisor of* n *and let* σ *be a fixed divisor of* $n/(k+1)$. *Moreover, let* $k+1$ *be a divisor of* $p^\sigma - 1$. *Any multiple of* σ *in the set* $0, 1, 2, \cdots, n-1$ *can be written in just one way in the form* $\{(k+1)s+\alpha\}\sigma$ *where* $0 \leqq \alpha \leqq k$ *and* s *is a non-negative integer. For every such multiple of* σ *form the entire set of homogeneous transformations*

$$(\bar{\text{B}}) \qquad \begin{aligned} \rho x_0' &= \beta_{s\alpha} x_0^{p^{\{(k+1)s+\alpha\}\sigma}}, && (\beta_{s\alpha} \neq 0), \\ \rho x_i' &= \sum_{j=0}^{k} \beta_{ijs\alpha} x_j^{p^{\{(k+1)s+\alpha\}\sigma}}, && (i = 1, 2, \cdots, k), \end{aligned}$$

in which each determinant of a transformation is equal to ω^α *times a* $(k+1)$*-th power in* $GF[p^n]$, ω *being a primitive mark of* $GF[p^n]$. *The totality of these transformations forms a subgroup* $EH_\sigma(k, p^n)$ *of* EC *which is of index* $(k+1)\sigma$ *in* EC. *Moreover,* EH_σ *is contained in* EC_σ *and is of index* $k+1$ *in* EC_σ. *The group* EH_σ *is generated by* $E\bar{C}_0$ *and the transformations of the form*

$$(\bar{\text{C}}) \qquad \rho x_i' = \omega^{t_i} x_i^{p^{\{(k+1)s+\alpha\}\sigma}}, \qquad (i = 0, 1, 2, \cdots, k),$$

where $t_0 + t_1 + t_2 + \cdots + t_k \equiv \alpha \bmod (k+1)$. *When considered as a permutation group on the* p^{kn} *points of* $EG(k, p^n)$, *the group* $EH_\sigma(k, p^n)$ *is doubly transitive.*

That the transformations named in the first paragraph of the theorem form a group is readily verified, as is also the fact that it is generated in the way indicated. It is also easily shown that EC_0 is invariant under transformation by the last collineation defined in the paragraph. As regards this

first paragraph of the theorem it remains to show that the order given for the group is correct. For this purpose we notice that a necessary and sufficient condition on the coefficients $\beta_{ij\tau}$ is that for each τ the determinant

$$\begin{vmatrix} \beta_{11\tau} & \beta_{12\tau} & \cdot & \cdot & \cdot & \beta_{1k\tau} \\ \beta_{21\tau} & \beta_{22\tau} & \cdot & \cdot & \cdot & \beta_{2k\tau} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \beta_{k1\tau} & \beta_{k2\tau} & \cdot & \cdot & \cdot & \beta_{kk\tau} \end{vmatrix}$$

shall be different from zero. The number of choices of these β's and β_τ satisfying this condition for fixed τ is known (compare theorem I of § 12) to be

$$\prod_{i=0}^{k-1} (p^{kn} - p^{in}).$$

The coefficients $\beta_{i0\tau}$ may each be chosen in p^n different ways for each value of τ; and hence the set for each value of τ may be chosen in p^{kn} different ways. Taking $\tau = 0$ we see that the number of transformations in EC_0 is the number given in the theorem. From this it follows readily that EC has the order stated.

The group EC_0 has been briefly treated by Veblen and Bussey (*l. c.*, p. 255). It is obviously equivalent to the general linear (non-homogeneous) group on k variables.

After this the proofs of the statements in the second and third paragraphs of the theorem are immediate.

To establish the transitivity properties named in the fourth paragraph note first that there is in $P(k, p^n)$ a transformation that carries any $k+2$ points of $PG(k, p^n)$, no $k+1$ of which are on the same $(k-1)$-space, into any like set of $k+2$ points and that in each of two such sets two points may be taken at will in $EG(k, p^n)$ while the remaining k points may be chosen from the $(k-1)$-space $x_0 = 0$. This transformation leaves invariant this $(k-1)$-space; hence it belongs to $EP(k, p^n)$. Hence $EP(k, p^n)$, considered as a permutation group on the points of $EG(k, p^n)$, is doubly transitive.

This transitivity property may also be established analytically and thus a verification may be had of the geometric property on which the previous proof is based. Let A and B be any two points of $EG(k, p^n)$. Then there is obviously a transformation in $EP(k, p^n)$ taking A into the point $(1, 0, 0, \cdot\cdot\cdot, 0)$. Let C be the point into which this transformation takes B. To establish the named transitivity property it is then sufficient to show that C may be taken, by a transformation of $EP(k, p^n)$, into $(1, 1, 0, 0, \cdot\cdot\cdot, 0)$ while $(1, 0, 0, \cdot\cdot\cdot, 0)$ remains invariant, or, what is equivalent, that

$(1, 1, 0, 0, \cdots, 0)$ may be so taken into the point C. The transformations which are available for this are those in which each β_{i00} is zero. Then the point $(1, 1, 0, 0, \cdots, 0)$ goes into the point $(\beta_0, \beta_{110}, \beta_{210}, \cdots, \beta_{k10})$. It is obvious that the β's may be chosen so that this is the point C. Hence the named transitivity property is established analytically.

It remains to treat further the case in which $k > 1$ and $p^n = 2$. For this purpose consider those transformations of $EP(k, 2)$ which leave fixed a given point P of $EG(k, 2)$. This group is obviously simply isomorphic with the projective group in $PG(k-1, 2)$, whence it may be seen that it is doubly transitive on the points of $EG(k, 2)$ exclusive of the point P. Hence $EP(k, 2)$ is triply transitive on the points of $EG(k, 2)$.

The remaining statement in the fourth paragraph of the theorem is now obviously true. Hence the part of the theorem which is contained in that paragraph is demonstrated.

To establish the transitivity properties named in the fifth paragraph of the theorem, let us denote any two points C and D of $EG(k, p^n)$ by

$$(\beta_\tau, \beta_{10\tau}, \beta_{20\tau}, \cdots, \beta_{k0\tau}) \text{ and } (\beta_\tau, \beta_{10\tau}+\beta_{11\tau}, \beta_{20\tau}+\beta_{21\tau}, \cdots, \beta_{k0\tau}+\beta_{k1\tau}), \quad (\beta_\tau \neq 0).$$

Since C and D are distinct by hypothesis it follows that at least one $\beta_{\mu 1\tau}$ is different from zero. Let λ be a fixed quantity such that $\beta_{\lambda 1\tau} \neq 0$. Then take $\beta_{\lambda s\tau} = 0$ when $s > 1$. Taking the quantities β, as thus defined, to be the coefficients in the transformation $(\bar{A})$ which are denoted by the same symbols, we see that the points $(1, 0, 0, \cdots, 0)$ and $(1, 1, 0, 0, \cdots, 0)$ are transformed by $(\bar{A})$ into C and D respectively. Now if $k > 1$ the remaining coefficients in the transformation can be so determined that the determinant of the transformation shall have any preassigned value. Hence these coefficients may be chosen so that the transformation belongs to the group $E\bar{C}_d(k, p^n)$. From this it follows that $E\bar{C}_d(k, p^n)$ is doubly transitive when $k > 1$. It is easy to treat analytically the case when $k = 1$ and to show that $E\bar{C}_d(1, 2^n)$ is doubly transitive while $E\bar{C}_d(1, p^n)$, for $p > 2$, is only singly transitive. To prove that this singly transitive group is primitive we observe that its elements may be denoted in non-homogeneous coördinates by the transformations $t' = \alpha t + \beta$ where α is a square in $GF[p^n]$ and β is any mark of $GF[p^n]$. Then it contains the transformation $t' = \omega^2 t$ where ω is a primitive mark of $GF[p^n]$. The corresponding permutation consists of two cycles each of order $\frac{1}{2}(p^n - 1)$. All the letters in either cycle must belong to the same set of imprimitivity if the group is imprimitive, whence it follows readily that the group is primitive.

[It may be remarked in passing that the set of transformations $t' = \alpha t + \beta$ where α runs over the λ-th powers in $GF[p^n]$ and β over all the marks of $GF[p^n]$, λ being a proper factor of $p^n - 1$, form a singly transitive group of degree p^n and order $(1/\lambda)p^n(p^n - 1)$; and that this set of groups contains other primitive groups than those named in the preceding paragraph. In particular, this group is primitive when λ is a factor of $p - 1$, as may be readily shown. There are also other conditions under which it may readily be proved that the group is primitive. There are also cases in which the group is imprimitive.]

It remains to prove the statements in the last paragraph of the theorem.

The fact that the transformations $(\bar{\mathrm{B}})$ form a group may be proved in the same way as the corresponding fact was established in the case of theorem I of § 12. The proof will therefore not be given. That EH_σ has the named generators is then proved in an obvious manner. That EH_σ has the named indexes in the groups mentioned is proved in the same way as that in which the corresponding results in theorem I of § 12 were established.

The transitivity property stated in the conclusion of the theorem may be established by the method employed in establishing the transitivity properties of $E\bar{C}_d(k, p^n)$. The proof is therefore omitted. The result for $k = 1$ is given by Mathieu (*l. c.*, p. 38).

This completes the proof of the theorem.

If the coefficients β_{i00} in $(\bar{\mathrm{A}})$, $i = 1, 2, \cdots, k$, are zero, then the point $(1, 0, 0, \cdots, 0)$ is left invariant by the transformation $(\bar{\mathrm{A}})$, and conversely. Hence we have an obvious analytical representation of that subgroup of each group in theorem I which consists of all the transformations in it which leave $(1, 0, 0, \cdots, 0)$ invariant. Moreover the transitivity properties of these subgroups follow immediately from the corresponding properties of the groups as given in the theorem. The subgroup of $EC_0(k, p^n)$ which leaves $(1, 0, 0, \cdots, 0)$ fixed is obviously equivalent to the general linear homogeneous group on k variables, as Veblen and Bussey have pointed out (*l. c.*, p. 255).

It is obvious that the group $EC(k, p^n)$ is multiply isomorphic with the group $C(k - 1, p^n)$ in the $PG(k - 1, p^n)$ defined by the equation $x_0 = 0$. By a comparison of the orders of these two groups it is then readily shown that the isomorphism is $p^{kn}(p^n - 1)$ to 1. In a transformation $(\bar{\mathrm{A}})$ of $EC(k, p^n)$ a variation in the coefficients β_τ, $\beta_{i0\tau}$ for $i = 1, 2, \cdots, k$ and τ fixed has no effect on the permutation in the $(k - 1)$-space $x_0 = 0$; and the variation of these coefficients gives $p^{kn}(p^n - 1)$ different transformations in $EC(k, p^n)$ corresponding to a given transformation in the subspace. Corresponding to the identity in $C(k - 1, p^n)$ we have therefore the $p^{kn}(p^n - 1)$ transformations

$$\rho x_0' = \beta_T x_0, \quad \rho x_i' = \beta_{i0T} x_0 + x_i, \qquad (i = 1, 2, \cdots, k),$$

in $EC(k, p^n)$. It is obvious that this carries the point $(1, 0, 0, \cdots, 0)$ to any assigned point in $EG(k, p^n)$, whence this subgroup is transitive in $EG(k, p^n)$. From this it follows that every subgroup of $EP(k, p^n)$ containing all the transformations of $EP(k, p^n)$ corresponding (in the named isomorphism) to a given subgroup of $P(k-1, p^n)$ is transitive. From this it follows that for every subgroup S of $C(k-1, p^n)$ there is a corresponding subgroup T of $EC(k, p^n)$, transitive on the p^{kn} points of $EG(k, p^n)$, the latter subgroup having with the former a $p^{kn}(p^n-1)$ to 1 isomorphism. Moreover, if the former subgroup is transitive the latter is doubly transitive, a fact which may be established as follows. The largest subgroup of T which leaves fixed one point A of $EG(k, p^n)$ contains a transformation carrying any line through A into any other line through A. Hence any given point in $EG(k, p^n)$, other than A, can be carried by a transformation of T into a point B of $EG(k, p^n)$ on any other line through A, while A itself remains fixed. Then, holding this latter line fixed, as well as the point A on it, we can take a transformation* in T which leaves point-wise invariant the subspace $x_0 = 0$ and carries B to any point C in $EG(k, p^n)$ and on the line AB. Hence the subgroup of T which leaves A fixed carries any given point of $EG(k, p^n)$ other than A to any such point. Hence the largest subgroup of T which leaves A fixed is transitive on the $p^{kn}-1$ points of $EG(k, p^n)$ other than A. Hence T itself is doubly transitive on the points of $EG(k, p^n)$. When S is intransitive it is easy to show in a similar way that T is only simply transitive.

We have thus demonstrated the following theorem, except for the statements about the primitivity of the singly transitive subgroups of $EC(k, p^n)$.

THEOREM II. *The group $EC(k, p^n)$ has a $p^{kn}(p^n-1)$ to 1 isomorphism with the group $C(k-1, p^n)$ on the points of the subspace $x_0 = 0$. The subgroup T of $EC(k, p^n)$ having a $p^{kn}(p^n-1)$ to 1 isomorphism with a given subgroup S of $C(k-1, p^n)$ and corresponding to it in the isomorphism just mentioned is a transitive group, when considered as a permutation group on the p^{kn} points of $EG(k, p^n)$. Moreover, when S is transitive, the group T is doubly transitive; otherwise it is simply transitive. When S is intransitive, a necessary and sufficient condition that the simply transitive group T is primitive is that it is generated by the largest subgroup leaving the point*

---

* If A is taken to be the point $(1, 0, 0, \cdots, 0)$, as it may without loss of generality, the available transformation is of the form

$$x_0' = x_0, \quad x_i' = \beta x_i, \qquad \beta \neq 0, \quad i = 1, 2, \cdots, k.$$

$(1, 0, 0, \cdots, 0)$ *fixed and any (every) single transformation whatever of* T *that does not leave this point fixed.*

It remains to prove the statement in the last sentence. It is an immediate consequence of the general theorem* that a necessary and sufficient condition that a transitive group G is imprimitive is that the largest subgroup of G which omits one letter is contained in a larger proper subgroup of G.

Every line in the Euclidean k-space $EG(k, p^n)$ has a point in common with the projective $(k-1)$-space $x_0 = 0$ which was excluded from $PG(k, p^n)$ in forming $EG(k, p^n)$. With a line of $EG(k, p^n)$ and a point of it not on this line we may form a Euclidean plane lying in $EG(k, p^n)$; as a plane of $PG(k, p^n)$ it contains a line in the excluded $(k-1)$-space. With such a plane and an additional point of $EG(k, p^n)$ we may form a three-space which is composed of a Euclidean three-space and a plane lying in the excluded $(k-1)$-space. It is clear that this process may be continued and that one may conclude to the existence in $EG(k, p^n)$ of a Euclidean m-space $EG(m, p^n)$ for every value m of the set $1, 2, \cdots, k-1$; and in each case the remainder of the projective space $PG(m, p^n)$ which contains $EG(k, p^n)$ lies in the excluded $(k-1)$-space $x_0 = 0$.

Now any collineation group in $EG(k, p^n)$ obviously permutes among themselves the m-spaces $EG(m, p^n)$ contained in $EG(k, p^n)$. Hence each of the named groups in theorems I and II, interpreted there as a permutation group on the points of $EG(k, p^n)$, may likewise be interpreted as a permutation group on the lines of $EG(k, p^n)$, or on its planes, or on its three-spaces, or in general on its m-spaces. The several permutation groups arising in this way from one and the same transformation group are obviously simply isomorphic each to each so that they are identical as abstract groups.

Hence we have the following theorem.

THEOREM III. *Any collineation group in* $EG(k, p^n)$ *may be interpreted as a permutation group on the included m-spaces* $EG(m, p^n)$ *for each value m of the set* $1, 2, \cdots, k-1$. *The several permutation groups, obtained by varying the value of m, are simply isomorphic each to each.*

We shall next prove the following theorem.

THEOREM IV. *Let T and S have the same meanings as in theorem II. If S is transitive on the points of the* $(k-1)$*-space* $x_0 = 0$, *then, the group T is transitive when interpreted as a permutation group on the lines of* $EG(k, p^n)$. *If S is transitive on the projective l-spaces contained in the projective* $(k-1)$-

* See Miller, Blichfeldt and Dickson's *Theory of Finite Groups*, p. 39.

space $x_0 = 0$, *then* T *is transitive on the Euclidean* $(l+1)$*-spaces contained in* $EG(k, p^n)$; *this group* T *is imprimitive.*

The truth of the statement contained in the second sentence of the theorem is an obvious consequence of theorems II and III. To prove the statement in the last sentence we observe first that T contains a transformation carrying one point of $EG(k, p^n)$ into any other while at the same time the projective $(k-1)$-space is left pointwise invariant. Now any Euclidean $(l+1)$-space in $EG(k, p^n)$ may be defined by a projective l-space in the subspace $x_0 = 0$ and a point of $EG(k, p^n)$, it being understood that all points of $EG(k, p^n)$ collinear with the given point and the given l-space constitute the named $(l+1)$-space. Now let A and B be two Euclidean $(l+1)$-spaces so defined and let P and Q be the points in $EG(k, p^n)$ used in thus defining them. Leaving the subspace $x_0 = 0$ pointwise invariant, take P to Q by means of a transformation belonging to T. Then holding Q fixed, take the l-space of A which is in the subspace $x_0 = 0$ into the corresponding l-space of B by means of an element of T. These two transformations taken in order carry A into B. Hence T has the required property of transitivity.

It remains to be shown that the group T is imprimitive on the named $(l+1)$-spaces. For this purpose it is sufficient to observe that all the $(l+1)$-spaces of $EG(k, p^n)$ which are based, in the way indicated, on a given l-space of the subspace $x_0 = 0$ are permuted among themselves when that l-space is left invariant and that they are transformed into a like set of $(l+1)$-spaces when the given l-space is transformed into another like l-space.

14. *Collineation Groups Leaving Other Subspaces Invariant.*

We shall now prove the following theorem:

THEOREM. *The group* $C^{(l)}(k, p^n)$ *consisting of all transformations of the form*

$$\text{(A)} \qquad \begin{aligned} \rho x_i' &= \sum_{j=0}^{l} \beta_{ij\tau} x_j^{p^\tau}, \qquad && (i = 0, 1, 2, \cdots, l), \\ \rho x_i' &= \sum_{j=0}^{k} \beta_{ij\tau} x_j^{p^\tau}, \qquad && (i = l+1, l+2, \cdots, k), \end{aligned}$$

where $0 \leqq l < k$, τ *runs over the sequence* $0, 1, 2, \cdots, n-1$, *and the coefficients* β *are marks of* $GF[p^n]$, *is a collineation group in* $PG(k, p^n)$ *which leaves invariant the subspace* $PG(k-l-1, p^n)$ *defined by the equations*

$$x_0 = 0, \; x_1 = 0, \cdots, x_l = 0.$$

It also leaves invariant the complementary set of $p^{(k-l)n}+p^{(k-l+1)n}+\cdots+p^{kn}$ *points in* $PG(k,p^n)$. *Its order is*

$$\frac{np^{(k-l)(l+1)n}}{p^n-1}\cdot\prod_{i=0}^{l}(p^{(l+1)n}-p^{in})\cdot\prod_{i=0}^{k-l-1}(p^{(k-l)n}-p^{in}).$$

The group is generated by its subgroup $C_0^{(l)}(k,p^n)$ *for which* $\tau=0$ *and the collineation*

$$\text{(B)}\qquad \rho x_i'=x_i^p,\qquad (i=0,1,2,\cdots,k).$$

The last element transforms $C_0^{(l)}(k,p^n)$ *into itself.*

For each proper divisor d *of* n *the group* $C^{(l)}(k,p^n)$ *has an obvious subgroup* $C_d^{(l)}(k,p^n)$ *of index* d *generated by* $C_0^{(l)}(k,p^n)$ *and the* d*-th power of the collineation* (B). *Moreover* $C_d^{(l)}(k,p^n)$ *has an obvious subgroup* $\bar{C}_d^{(l)}(k,p^n)$ *of index* μ *consisting of those transformations of* $C_d^{(l)}(k,p^n)$ *whose determinants are* $(k+1)$*-th powers,* μ *being the greatest common divisor of* $k+1$ *and* p^n-1.

The common subgroup of $C^{(l)}(k,p^n)$ *and the group* $H_\sigma(k,p^n)$ *of theorem I of* § 12 *consists of the entire set of transformations of the form*

$$\rho x_i'=\sum_{j=0}^{l}\beta_{ijsa}x_j^{p^{\{(k+1)s+a\}\sigma}},\qquad (i=0,1,2,\cdots,l),$$
$$\rho x_i'=\sum_{j=0}^{k}\beta_{ijsa}x_j^{p^{\{(k+1)s+a\}\sigma}},\qquad (i=l+1,l+2,\cdots,k),$$

the notation being that of theorem I of § 12 *and the determinant* $|\beta_{ijsa}|$ *being restricted as in that theorem.*

The group $C_0^{(l)}(k,p^n)$ *is transitive when interpreted as a permutation group on the set of* $p^{(k-l)n}+\cdots+p^{kn}$ *points mentioned in the first paragraph of the theorem.*

That the given set of transformations form a group leaving invariant the named subspace, and hence the complementary set of points, is obvious. To determine the order of the group we notice first that the determinant of the coefficients $\beta_{ij\tau}$ for i and j running over the set $0,1,2,\cdots,l$ must be different from zero; whence it follows (from a comparison with theorem I of § 12) that these coefficients can be chosen in

$$\prod_{i=0}^{l}(p^{(l+1)n}-p^{in})$$

different ways, τ remaining fixed. The coefficients $\beta_{ij\tau}$ for i and j running over the set $l+1,\ l+2,\cdots,k$ and τ remaining fixed can then be chosen

independently in any way so that their determinant shall be different from zero; and hence they can be chosen in

$$\prod_{i=0}^{k-l-1} (p^{(k-l)n} - p^{in})$$

different ways. Then for τ still fixed each of the remaining $(k-l)(l+1)$ coefficients β can be chosen independently in p^n ways, so that altogether this set of coefficients can be chosen in

$$p^{(k-l)(l+1)n}$$

different ways. Finally there are n values for τ. Hence the order of the group is the product of n and the three numbers just determined, all divided by $p^n - 1$, this divisor being introduced to allow for the factor of proportionality. From this it follows that the order of the group is that stated in the theorem.

That the group is generated in the way indicated is obvious.

The propositions in the second paragraph of the theorem are obvious in view of the corresponding parts of theorem I of § 13.

The proposition in the third paragraph of the theorem has an obvious demonstration in view of the proof of the corresponding part of theorem I of § 12.

Since any $k+2$ points no $k+1$ of which are on a $(k-1)$-space can be carried by the projective group into any other such set, it is obvious that an l-space may be held fiexd while any point not on it is transformed into any other such point. Thence follows readily the truth of the last proposition in the theorem.

Grundlagen der kombinatorischen Logik.

TEIL II.*

von H. B. CURRY.

C. DARSTELLUNG DER KOMBINATIONEN DURCH KOMBINATOREN IN DER NORMALFORM.

In diesem Abschnitte gebrauchen wir gewisse Zeichen, die wir Variablen nennen wollen. Diese Variablen sind nur ein Hilfsmittel, womit wir zeigen können, dass eine gewisse Art von Vollständigkeit und Verträglichkeit des Grundgerüstes vorliegt. Sie sind nicht als Ableitungen des Grundgerüstes anzusehen. Die Ausführungen dieses Abschnitts haben daher mit der formalen Darstellung nichts zu tun, sondern sie betreffen die Verwandtschaft zwischen dieser und der gewöhnlichen Logik. Diese Variablen sind als Etwase ohne besondere Eigenschaften zu behandeln.

Die Hauptergebnisse dieses Abschnitts sind die letzten Sätze von § 1 und § 5. Unter den ersten kommt der Hauptsatz I von Abschnitt A vor; dagegen macht § 5, Satz 2 den Kern des Hauptsatzes II aus.

§ 1. *Allgemeines über Reduktion und Entsprechen; ihre Eindeutigkeit.*

Festsetzung 1. In dem Folgenden betrachten wir Ausdrücke, die aus gewissen Variablen $x_1, x_2, x_3, \cdots$ und Etwase formal aufgebaut werden, d. h. so dass die Variablen als Etwase ohne besondere Eigenschaften behandelt werden. Auf solche Ausdrücke werden die vorhergehenden Festsetzungen und Definitionen ausgedehnt.

Festsetzung 2. Wir betrachten nun ein X, das eine Kombination von Kombinatoren und Variablen ist. Wir nehmen an, dass X mit den nach I C, Def. 1, erlaubten Auslassungen von Klammern geschrieben ist, und dass alle die anderen in den vorigen Abschnitten definierten Bezeichnungen durch ihre Definitionen ersetzt sind. Dann ist X von der Form $(X_0X_1X_2 \cdots X_n)$, wo X_0 entweder B, C, W, K oder eine Variable x_k ist, und die X_i für $i > 0$ Ausdrücke von derselben Form wie X sind.

Inbezug auf einen solchen X setzen wir zwei Arten von Reduktionsprozessen fest, wie folgt:

1.) Wenn X_0, B, C, W, oder K ist und n gross genug ist, so dürfen wir für

---

* Teil I erschien in diesem Journal, Bd. 52 (1930), S. 509-536.

$X_0X_1X_2$ bzw. $X_0X_1X_2X_3$ sein Äquivalent nach der betreffenden Regel B, C, W oder K ersetzen, z. B., wenn $X_0 \equiv B$,* so haben wir $X_1(X_2X_3)X_4 \cdots X_n$ anstatt $X_0X_1X_2X_3X_4 \cdots X_n$. Eine solche Ersetzung soll ein *Reduktionsprozess erster Art* heissen.

2.) Es mag sein, dass ein Bestandteil von X (d. h. ein eingeklammerter in X erscheinender Ausdruck) durch einen Reduktionsprozess erster Art umgeformt werden kann. Eine solche Umformung soll ein *Reduktionsprozess zweiter Art* heissen, wenn ein Reduktionsprozess erster Art sowohl für den Gesamtausdruck wie auch für jeden Teilausdruck, der den betreffenden einschliesst oder links von ihm steht, unmöglich ist.

Festsetzung 3. Ein Ausdruck X *reduziert* sich auf einen anderen Y, wenn durch Anwendung dieser Prozesse X in Y umgeformt wird, und zwar *im ersten Sinne,* wenn nur Prozesse erster Art nötig sind, und im *zweiten Sinne,* wenn auch Prozesse zweiter Art nötig sind. Dass X sich auf Y reduziert wird auch durch das Zeichen $X \doteq Y$ ausgedruckt.

Festsetzung 4. Es sei ein Ausdruck X_m gegeben, der x_m, aber keine x_n, $n > m$, enthält, und der ferner nicht von der Form $X_{m-1}x_m$ ist. Dann denken wir an die unendliche Zeichenfolge, welche entsteht, wenn man rechts von X_m die Variablen $x_{m+1}, x_{m+2}, \cdots$ ad infin. setzt; diese heisst die durch X_m bestimmte *Folge.* Der Teil dieser Folge, welcher einem bestimmten x_n, $n > m$, vorangeht, ist ein Ausdruck, der ein *Abschnitt* der Folge heisst. Also ist X_m selbst ein Abschnitt der durch ihn bestimmten Folge.

Festsetzung 5. Ein Ausdruck X enthält eine Variable x_m *wesentlich,* wenn für $n >$ den Index irgendeiner in X erscheinenden Variablen, und für $p \geqq 0$, in der Reduktion von $(Xx_nx_{n+1} \cdots x_{n+p})$ die Variable x_m nie ausfällt.† Z. B. der Ausdruck $B(Kx_1)x_2$ enthält x_1, wesentlich, aber nicht x_2. Die höchste wesentlich erscheinende Variable in X heisst der *Grad* von X. (Wenn keine Variable wesentlich erscheint, so heisst der Grad 0).

Festsetzung 6. In der Reduktion eines Ausdrucks X auf einen anderen Y heisst eine Variable x_n *nicht gestört,* wenn 1) X von der Form $X'x_nx_{n+1} \cdots x_{n+p}$ ist, wo X' die Variablen $x_n, x_{n+1}, \cdots, x_{n+p}$ nicht wesentlich enthält, 2) Y von einer ähnlichen Form $Y'x_nx_{n+1} \cdots x_{n+p}$ ist, 3) X' sich auf Y' reduzieren lässt. Sonst heisst eine in X erscheinende Variable gestört.

Festsetzung 7. Ein Ausdruck X *entspricht* einer Folge $\mathfrak{X}$, wenn die fol-

* Der leser soll bemerken, dass Ausdrücke wie $X \equiv Y$ und $\vdash X$. Sätze bedeuten. (s. IC).

† Natürlich soll x_m nicht in X selbst fehlen.

gende Bedingung erfüllt ist: es gibt ein $p \geqq 0$, so dass der Ausdruck $(X x_{n+1} x_{n+2} \cdots x_{n+p})$, wo n der Grad von X ist, sich auf einen Abschnitt von $\mathfrak{X}$ reduziert, und zwar so, dass x_{n+p}, wenn $n + p > 0$ ist, nicht ausgelassen wird. Wenn x_{n+q} die höchste in dieser Reduktion gestörte Variable ist (bzw. $q = 0$, wenn keine nicht in X wesentlich erscheinende Variable gestört wird), so sagen wir, dass X der Folge *mit der Ordnung* q entspricht.* Endlich sprechen wir von einem Entsprechen *im ersten bzw. zweiten Sinne,* wenn die Reduktion sich im ersten bzw. zweiten Sinne vollzieht.

Festsetzung 8. Zwei Ausdrücke X und Y heissen *äquivalent* im

1) *ersten Sinne,* wenn sie denselben Grad haben und derselben Folge von lauter Variablen entsprechen,

2) *zweiten Sinne,* wenn sie denselben Grad haben, und derselben Folge von lauter Variablen mit derselben Ordnung entsprechen,

3) *dritten Sinne,* wenn sie denselben Grad haben, und derselben Folge von lauter Variablen in demselben Sinne entsprechen.

4) *vierten Sinne,* wenn sie denselben Grad haben, und derselben Folge von lauter Variablen in demselben Sinne und mit derselben Ordnung entsprechen.

Bemerkung: Die folgenden Sätze haben als Zweck den Beweis, dass, wenn eine Formel der Form $\vdash X = Y$ aus unserem Grundgerüst ableitbar ist, ein gewisser Sinn von Äquivalenz zwischen X und Y besteht. Eine gewisse Art von Übereinstimmung mit Logik und Unabhängigkeit wird dabei für die kombinatorischen Axiome gewährleistet. Dies ist die einzige solche Untersuchung dieser Abhandlung; eine allgemeine Vollständigkeits-, Widerspruchslosigkeits- oder Unabhängigkeitsuntersuchung wird von dieser Abhandlung ausgeschlossen.

Hilfssätze. Das Reduzieren ist seiner Definition nach ein eindeutiger Prozess, also haben wir leicht die folgenden Hilfsätze.

1. Wenn ein Ausdruck sich auf zwei verschiedene Ausdrücke reduziert, so reduziert einer dieser beiden sich auf den anderen.

---

* Man darf hier annehmen dass entweder $p = q$ oder $p = q + 1$ ist. Denn nach den Voraussetzungen reduziert $X x_{n+1} x_{n+2} \cdots x_{n+p}$ sich auf ein $\mathfrak{C} x_{n+q+1} x_{n+q+2} \cdots x_{n+p}$ und zwar so, dass $x_{n+q+1}, x_{n+q+2}, \cdots, x_{n+p}$ dabei ungestört werden. Daher reduziert $X x_{n+1} x_{n+2} \cdots x_{n+q+1}$ sich auf $\mathfrak{C} x_{n+q+1}$; und dieser ist ein Abschnitt der Folge $\mathfrak{F}$, näm., $\mathfrak{C} x_{n+q+1} x_{n+q+2} \cdots$. Wir wissen ja auch, dass $X x_{n+1} \cdots x_{n+q}$ sich auf $\mathfrak{C}$ reduziert, aber davon können wir nicht schliessen, dass immer $p = q$ sein kann, weil $\mathfrak{C}$ nicht ein Abschnitt der Folge $\mathfrak{F}$ ist, falls x_{n+q} in der Reduction ausfällt.

2. Ein Ausdruck kann nie auf zwei verschiedene Kombinationen von Variablen reduziert werden.

3. Ein Ausdruck kann nie zwei verschiedenen Folgen von Variablen entsprechen.

4. Wenn X und Y denselben Grad n haben, und wenn ferner die zwei Ausdrücke $(Xx_{n+1}x_{n+2}\cdots x_{n+p})$ und $(Yx_{n+1}x_{n+2}\cdots x_{n+p})$ sich auf dieselbe Kombination lauter Variablen reduzieren, so sind X und Y äquivalent in dem ersten Sinne.

5. Wenn X und Y denselben Grad n haben, und wenn ferner für jedes p, wofür einer der beiden Ausdrucke $(Xx_{n+1}x_{n+2}\cdots x_{n+p})$ und $(Yx_{n+1}x_{n+2}\cdots x_{n+p})$ auf eine Kombination von lauter Variablen reduziert wird, die beiden sich auf dieselbe Kombination reduzieren, so sind X und Y in dem zweiten Sinne äquivalent.

Satz 1. *Sind $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N, X, Y,$ Kombinationen von Kombinatoren und Variablen $x_1, x_2, \cdots, x_m$ derart, dass*

1) *für jedes i $(i = 1, 2, \cdots N)$ die beiden Ausdrücke $\mathfrak{A}_i$ und $\mathfrak{B}_i$ denselben Grad haben, und weiter derselben Folge mit derselben Ordnung und im ersten Sinne entsprechen,*

2) *X einer Folge von lauter Variablen entspricht,*

3) *aus den Voraussetzungen*

$$\vdash \mathfrak{A}_i = \mathfrak{B}_i \tag{1}$$

mit Benutzung nur der Eigenschaften der Gleichheit $(I\,D)$ folgt, dass

$$\vdash X = Y; \tag{2}$$

dann sind X und Y äquivalent im zweiten Sinne, und zwar, wenn jedes $\mathfrak{A}_i$ und jedes $\mathfrak{B}_i$ wirklich Kombinatoren enthält, im vierten Sinne.

Beweis: Es genügt, den Satz für den Fall zu beweisen, dass X aus Y durch eine einzige Ersetzung entsteht, nämlich der Ersetzung eines in X erscheinenden $\mathfrak{A}$ durch seinen Gegenwert $\mathfrak{B}$, oder umgekehrt. Das allgemeinste Y ergibt sich aus X durch eine Reihe von solchen Ersetzungen.

Nach Hp. 2 gibt es ein n, wofür $(Xx_{m+1}x_{m+2}\cdots x_{m+n})$ sich auf eine Kombination von $x_1, x_2, \cdots, x_{m+n}$ reduziert. Ich möchte diese Kombination Z nennen, und die Ausdrücke $(Xx_{m+1}x_{m+2}\cdots x_{m+n})$ bzw. $(Yx_{m+1}x_{m+2}\cdots x_{m+n})$ mit X' und Y' abkürzen. Ich zeige zunächst, dass Y' auf Z reduziert wird, und zwar, wenn die $\mathfrak{A}_i$ und $\mathfrak{B}_i$ alle wirklich Kombinatoren enthalten, in demselben Sinne.

$\mathfrak{A}$ sei der ersetzte Ausdruck in X und $\mathfrak{B}$ sein Gegenwert, so lässt Y' sich von X' nur dadurch unterscheiden, dass in Y' $\mathfrak{B}$ die Stelle von $\mathfrak{A}$ einnimmt. Dann können im Laufe der Reduktion die folgenden drei Möglichkeiten geschehen:

I. Wir kommen zu einer Form an, worin $\mathfrak{A}$ am Anfang steht, d. h. zu einer Form

$$(3) \qquad (\mathfrak{A}X_1X_2\cdots X_p),$$

wo die $X_1, X_2, \cdots, X_p$ Kombinationen von Kombinatoren und Variablen sind.

II. Ein eingeklammerter Teilausdruck, der $\mathfrak{A}$ enthält (bzw. $\mathfrak{A}$ selbst), wird als ein Ganzes durch K ausgestrichen.

III. $\mathfrak{A}$ bleibt innerhalb des Gesamtausdrucks (d. h. nicht am Anfang), bis in der Reduktion durch Prozesse zweiter Art die Reihe an es kommt, und dann steht es am Anfang eines Teilausdrucks der Form (3), wo $p \geqq 0$ ist.

Diese drei Möglichkeiten sind erschöpfend, weil Reduktion so definiert ist, dass $\mathfrak{A}$ sonst ein untrennbares Ganzes ist. Ich behandle die drei Fälle jetzt besonders.

Fall I. Nach der Voraussetzung dieses Falles reduziert X' sich auf einen Ausdruck X'' der Form (3). Dann reduziert sich Y' durch genau dieselbe Reihe von Reduktionsprozessen auf ein Y'' der Form

$$(4) \qquad (\mathfrak{B}X_1X_2\cdots X_p),$$

wo die $X_1, X_2, \cdots, X_p$ dieselben Ausdrücke wie in X'' sind.

Es werde nun angenommen, der Ausdruck $\mathfrak{A}' \equiv (\mathfrak{A}x_{m+1}x_{m+2}\cdots x_{m+p})$ reduziert sich im ersten Sinne auf einen Ausdruck $\mathfrak{C}$. Dann, wenn wir überall in dieser Reduktion $x_{m+1}, x_{m+2}, \cdots x_{m+p}$ durch $X_1, X_2, \cdots X_p$ ersetzen, liefert die so entstehende Folge von Ausdrücken wieder eine Reduktion im ersten Sinne. Daher reduziert sich X'' durch Prozesse erster Art auf ein X''', welches entsteht, wenn man in $\mathfrak{C}$ die betreffenden Einsetzungen macht. Eine ähnliche Bemerkung bezieht sich auf Y''.

Nach Hp. 1 entsprechen $\mathfrak{A}$ und $\mathfrak{B}$ beide derselben Folge $\mathfrak{F}$. r sei die Ordnung, womit $\mathfrak{A}$ dem $\mathfrak{F}$ entspricht. Dann zeige ich, dass $r \leqq p$ ist. In der Tat nehmen wir das Gegenteil an. Dann reduziert der Ausdruck $(\mathfrak{A}'x_{m+p+1}\ x_{m+p+2}\cdots x_{m+r+1})$ sich auf einen Abschnitt von $\mathfrak{F}$ und zwar so, dass x_{m+p+1} gestört wird.* Unter der durch diese Reduktion erzeugten Reihe von Ausdrücken gibt es ein $(\mathfrak{C}x_{m+p+1}x_{m+p+2}\cdots x_{m+r+1})$ derart, dass die Reduktion sich bis auf diesen Ausdruck ohne Störung von $x_{m+p+1}\cdots x_{m+r+1}$ erstrickt,

* S. Festsetzung 6, Anmerkung.

während im nächsten Schritte der Reduktion x_{m+p+1} gestört wird. Also muss $\mathfrak{C}$ von der Form

$$(5) \qquad (X_0''X_1''X_2'' \cdots X_q'')$$

sein, wo entweder 1) X_0'' B oder C ist und $q < 3$ ist, oder 2) X_0'' W oder K ist und $q < 2$ ist. Nach Festsetzung 6 reduziert $\mathfrak{A}'$ sich auf dieses $\mathfrak{C}$. Dann reduziert sich X'' nach dem vorigen Absatz auf ein X''' derselben Form (5). Aber in der weiteren Reduktion eines solchen X''' könnte der Kombinator X_0'' nie verschwinden, was der Voraussetzung, dass X' sich auf Z reduziert, widerspricht.

Also gilt $r \leqq p$. Dann reduzieren sich die beiden Ausdrücke $(\mathfrak{A}'x_{m+p+1})$ und $(\mathfrak{B}'x_{m+p+1})$ auf einen Abschnitt $(\mathfrak{C}x_{m+p+1})$ von $\mathfrak{F}$, und zwar so, dass x_{m+p+1} ungestört wird. Daher reduzieren sich $\mathfrak{A}'$ und $\mathfrak{B}'$ beide auf dasselbe $\mathfrak{C}$ (Festsetzung 6). Diese Reduktion geschieht weiterhin im ersten Sinne. Nach dem vorletzten Absatz reduzieren sich dann X'' und Y'' auf ein gemeinsames X''', und zwar im ersten Sinne. Weil X' auf Z reduziert wird, so reduziert sich X''', und also Y' auf Z. Weil die einzigen Reduktionsprozesse, die in den Reduktionen von X' und Y' verschieden sind, zu der ersten Art gehören, so reduzieren X' und Y' sich auf Z in demselben Sinne.

Fall II. Durch eine Reihe von Reduktionsprozessen reduziert X' sich auf einen Ausdruck X'', der einen Teilausdruck der Form $(KX_1X_2 \cdots X_p)$ enthält, wo $\mathfrak{A}$ in X_2 enthalten ist, und zwar so, dass beim nächsten Schritte die Reduktion auf einen X'' führt, der sich vom X'' nur dadurch unterscheidet, dass der obige Teilausdruck durch $(X\, X_3 \cdots X_p)$ ersetzt ist. Genau dieselbe Reihe von Prozessen reduziert Y' auf einen Ausdruck Y'', der sich von X'' nur darin unterscheidet, dass $\mathfrak{A}$ die Stelle von $\mathfrak{B}$ einnimmt. Beim nächsten Schritte, der derselbe Prozess wie im vorigen Falle ist, kommen wir wieder auf X'''. Daher reduzieren X' und Y' sich durch dieselbe Reihe von reduktionsprozessen auf denselben Ausdruck. Infolgedessen reduzieren sie sich endlich auf dieselben Kombination, und zwar, weil die beiden Reihen von Prozessen dieselben sind, in demselben Sinne.

Fall III. Nach der Voraussetzung reduziert X' sich auf einen Ausdruck X'', der einen Teilausdruck der Form (3) enthält, und zwar so, dass die weitere Reduktion von X'' durch die Reduktion dieses Teilausdrucks fortgesetzt wird. Dann reduziert Y' sich auf ein Y'', welches sich von X'' nur darin unterscheidet, dass der Ausdruck (4) anstatt (3) erscheint.

Weil die Bedingungen von Fall I für diese Teilausdrücke (3) und (4) erfüllt sind, so reduzieren diese Teilausdrücke sich auf dieselben Kombinationen. Weil X'' und Y'' sonst identisch sind, so reduzieren X'' und Y'',

und daher auch X' und Y' sich auf denselben Ausdruck. Infolgedessen werden X' und Y' auf dieselbe Kombination von Variablen reduziert.

Wenn $\mathfrak{A}$ und $\mathfrak{B}$ wirklich Kombinatoren enthalten, so sind Reduktionsprozesse zweiter Art in den beiden Fällen erforderlich. Deshalb werden sie auf diese Kombination in demselben Sinne reduziert.

Es ist nun bewiesen, dass X' und Y' sich auf dasselbe Z reduzieren. Daraus folgt zunächst, dass X und Y denselben Grad haben; denn jede Variable, die in der Reduktion von X' verschwindet, verschwindet auch in der Reduktion von Y', und umgekehrt. Dieser Grad sei dann μ. Setzen wir in den obigen Beweis $x_{\mu+j}$ statt x_{m+j} ein, so folgt, dass die neuen X' und Y' auch auf eine gemeinsame Kombination lauter Variablen reduziert werden, wenn nur eines von den beiden sich auf eine solche Kombination reduziert. Also entsprechen X und Y derselben Folge mit derselben Ordnung (Hilfsatz 5), und auch, wenn die $\mathfrak{A}$ und $\mathfrak{B}$ wirklich Kombinatoren enthalten, in demselben Sinne, w. z. b. w.

SATZ 2. *Sind* $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N, X, Y,$ *Kombinationen von Kombinatoren und Variablen* $x_0, x_1, \cdots, x_m$ *derart, dass die Bedingungen von Satz 1 erfüllt sind, ausser dass in Hp. 3 bei der Ableitung von* (2) *aus* (1) *auch Benutzung von den Regeln B, C, W und K erlaubt wird; dann sind X und Y im zweiten Sinne äquivalent.*

Beweis: Wir können von X zu Y durch eine Reihe von Schritten übergehen, wovon jeder daraus besteht, dass wir entweder eine einzige Ersetzung aus den Formeln (1) machen, oder auch eine Regel B, C, W oder K einmal anwenden. Weiter dürfen wir unter einer solchen Anwendung den folgenden Prozess verstehen: zunächst setzen wir in einer Regel (B, C, W oder K) für die X, Y (und Z, wenn es erscheint) besondere Ausdrücke ein, sodass eine Formel $\mathfrak{A} = \mathfrak{B}$ entsteht, und dann machen wir in einem schon aus X abgeleiteten Ausdruck eine Ersetzung von $\mathfrak{A}$ durch $\mathfrak{B}$ oder umgekehrt.

Jetzt betrachten wir alle die Formeln, die in diese Weise aus allen den im Uebergang von X zu Y benutzten Anwendungen der betreffenden Regeln entstehen. Fügen wir diese Formeln zu den Formeln (1) hinzu. Dann sind alle die Bedingungen von Satz 1 für die erweiterte $\mathfrak{A}_1, \mathfrak{A}_2, \cdots \mathfrak{A}_N, \mathfrak{A}_{n+1}, \cdots, \mathfrak{A}_M, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N, \mathfrak{B}_{N+1}, \cdots, \mathfrak{B}_M, X, Y$ erfüllt. Also folgt der Satz aus Satz 1.

Es soll bemerkt werden, dass die Nebenbedingung für den strengen Satz 1, nämlich dass alle die $\mathfrak{A}_i$ und $\mathfrak{B}_i$ wirklich Kombinatoren enthalten, für die neue $\mathfrak{A}_{N+j}$ oder $\mathfrak{B}_{N+j}$ versagen mag, sogar wenn sie für die ursprüngliche $\mathfrak{A}_i$ und $\mathfrak{B}_i$ erfüllt ist.

Satz 3. *Wenn $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N$, X, Y Kombinatoren sind, die die Hypothesen von Satz 2 erfüllen; dann sind X und Y äquivalent im vierten Sinne.*

Beweis: Die $\mathfrak{A}_i$ und $\mathfrak{B}_i$, die sowohl in den ursprünglichen Formeln (1), als auch in denen, die dazu durch die Prozesse des Beweises von Satz 2 hinzugefügt werden, erscheinen, sind Kombinatoren und enthalten daher Kombinatoren. Also folgt der Satz aus Satz 1.

Satz 4. *Sind X, Y Kombinatoren, wofür*

1) *es folgt aus den transmutativen Axiomen mit Benutzung der Regeln B, C, W, K und den Eigenschaften der Gleichheit, dass $\vdash X = Y$,*

2) *mindestens einer der beiden einer Folge von lauter Variablen entspricht;*

dann sind X und Y äquivalent im vierten Sinne.

Beweis: folgt aus Satz 3, weil die betreffenden Axiome die Bedingungen der Formeln (1) erfüllen.

Satz 5. *Wenn $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_N, \mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_N$, X, Y Kombinationen von Variablen und Kombinatoren sind, derart, dass*

1) *für jedes i $(i = 1, 2, \cdots, N)$ $\mathfrak{A}_i$ und $\mathfrak{B}_i$ denselben Grad haben und weiter derselben Folge mit derselben Ordnung entsprechen,*

2) *X einer Folge von lauter Variablen entspricht,*

3) *aus den Formeln*

$$\vdash \mathfrak{A}_i = \mathfrak{B}_i \qquad (1 = 1, 2, \cdots, N) \tag{1}$$

mit Benutzung der Regeln B, C, W, K und der Eigenschaften der Gleichheit folgt, dass

$$\vdash X = Y; \tag{2}$$

dann sind X und Y äquivalent im zweiten Sinne.

Beweis: Zunächst sehen wir sofort, dass der Satz, wenn er für den Fall bewiesen ist, dass im Hp. 3) die Benutzung nur von den Eigenschaften der Gleichheit erlaubt ist, im allgemeinen durch das Verfahren, das ich in dem Beweis von Satz 2 benutzt habe, bewiesen werden kann. Es genügt daher, den Satz für jenen Fall zu beweisen.

Der Beweis verläuft nun wie der von Satz 1. Wir setzen ohne Beschränkung der Allgemeinheit voraus, dass Y sich aus X durch eine einzige Einsetzung, die von $\mathfrak{A}$ statt $\mathfrak{B}$, ergibt. Wir definieren X', Y' und Z wie dort, und schliessen, wie folgt, dass Y' auf Z reduziert wird. Wir unterscheiden dieselben drei Fälle wie im Satz 1.

Fall I. X' und Y' reduzieren sich auf X'' bzw. Y'' von der Form (3) bzw. (4).

Es sei nun angenommen, der Ausdruck $\mathfrak{A}'$ (definiert wie im Satz 1) reduziert sich auf einen Ausdruck $\mathfrak{C}$; dann, wenn wir überall in dieser Reduktion $x_{m+1}, x_{m+2}, \cdots, x_{m+p}$ durch $X_1, X_2, \cdots, X_p$ ersetzen, so schaffen wir eine Reihe von Ausdrücken, die, obgleich sie nicht immer eine Reduktion liefern müssen, doch nach Satz 2 (für $N=0$) immer zueinander im zweiten Sinne äquivalent sind. Infolgedessen muss X'', und daher auch X' mit einem X''', das aus $\mathfrak{C}$ durch die erwähnte Einsetzung entsteht, im zweiten Sinne äquivalent sein.

Nach Hp. 1 entsprechen $\mathfrak{A}$ und $\mathfrak{B}$ derselben Folge $\mathfrak{F}$. r sei die Ordnung, womit $\mathfrak{A}$ dem $\mathfrak{F}$ entspricht. Dann gilt $r \leqq p$. In der Tat sei angenommen, dass $r > p$ ist. Dann folgt, genau wie in Satz 1, dass $\mathfrak{A}'$ auf ein $\mathfrak{C}$ der Form (5) reduziert wird. Daher ist X'', nach dem vorigen Absatz, mit einem X''' der Form (5) im zweiten Sinne äquivalent. Dies ist aber unmöglich, weil X', und daher X'', einer mit Z anfangenden Folge lauter Variablen mit der Ordnung 0 entspricht, während X''' keiner Folge lauter Variablen mit der Ordnung 0 entsprechen kann.

Es folgt dann, wie im Satz 1, dass $\mathfrak{A}'$ und $\mathfrak{B}'$ sich auf dasselbe $\mathfrak{C}$ reduzieren. Daher sind X und Y nach dem vorletzten Absatz mit demselben X''' im zweiten Sinne äquivalent. Aber nach der Voraussetzung reduziert X' sich auf Z. Daher reduziert sich auch Y' auf Z.

Der Rest des Beweises verläuft genau wie im Satz 1.

SATZ 6. *Sind X und Y Kombinatoren, wofür*

1) *mindestens einer der beiden einer Folge von lauter Variablen entspricht,*

2) *aus den transmutativen und kommutativen Axiomen folgt, dass* $\vdash X = Y$;

dann sind X und Y äquivalent im zweiten Sinne.

Beweis: Folgt aus Satz 5, weil die betreffenden Axiome die Bedingungen der Formel (1) erfüllen.

SATZ 7. *Ax. I_2 ist nicht aus den übrigen kombinatorischen Axiomen mit Benutzung der Regeln B, C, K, W und den Eigenschaften der Gleichheit ableitbar.*

Beweis: Folgt aus Satz 6, weil die zwei Kombinatoren, die in Ax. I_2 auf den beiden Seiten des Zeichens = stehen, im dritten, aber nicht im zweiten Sinne äquivalent sind.

Satz 8. *Wenn wir in den Hypothesen von Sätzen 1-3 und 5 die folgenden Änderungen machen:*

1) $\mathfrak{A}_i$ *und* $\mathfrak{B}_i$ *brauchen nicht dieselbe Ordnung (inbezug auf ihr Entsprechen einer gemeinsamen Folge) zu haben,*

2) *nicht nur* X, *sondern auch* Y *einer Folge von lauter Variablen entspricht;*

dann folgen die Schlüsse dieser Sätze, wenn wir darin den vierten Sinn durch den dritten, und den zweiten Sinn durch den ersten ersetzen.

Beweis: Die einzigen Stellen in den Beweisen der betr. Sätze, wo wir die Voraussetzung über die Ordnung von den $\mathfrak{A}_i$ und $\mathfrak{B}_i$ benutzt haben, sind im Fall I unter den Sätzen 1 und 5, und zwar wird sie da nur benutzt, um zu beweisen, dass $\mathfrak{A}'$ und $\mathfrak{B}'$ sich auf ein gemeinsames $\mathfrak{C}$ reduzieren.

Diesen Schluss können wir auch im vorliegenden Falle erreichen. Es folgt ohne Benutzung der betr. Voraussetzung, dass entweder $\mathfrak{A}'$ und $\mathfrak{B}'$ sich auf ein gemeinsames $\mathfrak{C}$ reduzieren, oder einer der beiden auf einen Ausdruck der Form (5) reduziert wird. n sei nun so gewählt, dass nicht nur X', sondern auch Y' sich auf eine Kombination von lauter Variablen reduziert. Dies ist möglich nach Hp. 2 dieses Satzes. Dann folgt durch das Argument des dritten Absatzes des Falles 1 in den Sätzen 1 und 5, dass weder $\mathfrak{A}'$ noch $\mathfrak{B}'$ sich auf einen Ausdruck der Form (5) reduzieren lässt. Daher müssen sie sich auf einen gemeinsamen $\mathfrak{C}$ reduzieren.

Diese Änderung des n stört aber nichts in den Beweisen der betr. Sätze, ausser dass wir jetzt nicht schliessen können, dass X und Y dieselbe Ordnung haben. Also haben wir einen wirklichen Beweis, wenn wir die ganzen Beweise hindurch die Ersetzungen vom Schlusse dieses Satzes machen. Damit wird die Behauptung bewiesen.

Satz 9. *Wenn* X *und* Y *Kombinatoren sind, wofür*

1) *sowohl* X *wie auch* Y *einer Folge von lauter Variablen entspricht,*

2) *aus den transmutativen Axiomen und Ax.* I_2 *mit Benutzung der Eigenschaften der Identität und Regeln* B, C, K, W *folgt, dass* $\vdash X = Y$;

dann sind X *und* Y *im dritten Sinne äquivalent.*

Beweis: Folgt aus Sätzen 3 und 8.

Satz 10. *Die kommutativen Axiome sind nicht Folgerungen aus den anderen kombinatorischen Axiomen.*

Beweis: Die Kombinatoren, die in diesen Axiomen auf den beiden Seiten des Zeichens $=$ stehen, sind nicht im dritten Sinne äquivalent. Daher folgt der Satz aus Satz 9.

SATZ 11. *Sind X und Y Kombinationen von Variablen und Kombinatoren derart, dass*

1) *sowohl X wie auch Y einer Folge lauter Variablen entspricht,*

2) *aus den kombinatorischen Axiomen überhaupt mit Benutzung der Eigenschaften der Gleichheit und der Regeln B, C, K, W folgt, dass* $\vdash X = Y$;

dann haben X und Y denselben Grad, und sie entsprechen derselben Folge.

Beweis: Nach den Sätzen 5 und 8 sind X und Y im ersten Sinne äquivalent. Daher folgt der Satz gleich aus der Definition der Äquivalenz.

SATZ 12. *Sind X und Y Kombinationen lauter Variablen, wofür die Hp. 2 von Satz 11 erfüllt ist, so sind X und Y identisch.*

Beweis: Nach Satz 11 entsprechen X und Y derselben Folge; dies kann nur geschehen, wenn sie Abschnitte derselben Folge sind. Weiter haben sie nach Satz 11 denselben Grad; daraus folgt, dass sie genau derselbe Abschnitt sind.

Festsetzung 9. Ein Kombinator X stellt eine Kombination Y der Variablen $x_1, x_2, \cdots, x_n$ dann und nur dann dar, wenn aus den kombinatorischen Axiomen mit Benutzung der Eigenschaften der Gleichheit und der Regeln B, C, W, K folgt, dass

$$\vdash X x_1 x_2 \cdots x_n = Y.$$

SATZ 13. *Wenn ein Kombinator eine Kombination von $x_1, x_2, \cdots, x_n$ darstellt, so stellt er nur eine dar.*

Beweis: Folgt gleich aus Satz 12.

§ 2. *Normale Kombinationen und Folgen.*

Festsetzung 1. Unter einer *normalen Kombination* von $X_0, X_1, X_2, \cdots, X_n$ verstehen wir einen Ausdruck der Form

$$(X_0 Y_1 Y_2 \cdots Y_n),$$

wo jedes Y_i eine Kombination von $X_1, X_2, \cdots, X_n$ ist.

Festsetzung 2. Hiernach wird zuweilen auch das Zeichen x_0 als Variable gebraucht.*

* In der inhaltlichen Anwendung der vorliegenden Theorie wird im allgemeinen eine Funktion (wie ϕ in II A 3), die Stelle von x_0 einnehmen. Die Variable x_0 wird hiernach im allgemeinen nur für normalen Folgen usw. benutzt.

Festsetzung 3. Unter einer *normalen Folge* (von Variablen) verstehen wir eine Folge, die durch eine normale Kombination von $x_0, x_1, x_2, \cdots, x_n$ bestimmt ist (§ 1, Festsetzung 4), wo n irgendeine ganze Zahl > 0 ist. Solche normalen Folgen werden hiernach mit griechischen Buchstaben bezeichnet.

Festsetzung 4. Unter dem *Produkt* $(\eta \cdot \zeta)$ von zwei normalen Folgen η und ζ verstehen wir die folgendermassen bestimmte Reihe (von Variablen): Es sei

$$\eta = x_0 y_1 y_2 y_3 \cdots$$
$$\zeta = x_0 z_1 z_2 z_3 \cdots.$$

Ersetzt man dann in $z_1, z_2, \cdots$ die $x_1, x_2, x_3, \cdots$ bzw. durch $y_1, y_2, y_3, \cdots$, so ist das Resultat $(\eta \cdot \zeta)$.

SATZ 1. *Das Produkt von zwei normalen Folgen ist eine normale Folge.*

Beweis: η und ζ werden wie in der Festsetzung 4 bezeichnet und $(\eta \cdot \zeta)$ werde durch

$$x_0 u_1 u_2 u_3 \cdots$$

bezeichnet.

Nach der Definition einer Normalfolge gibt es ein m und ein n sodass 1) $(x_0 y_1 y_2 \cdots y_n)$ eine normale Kombination von $x_1, x_2, \cdots, x_m$ ist, die x_m wirklich enthält, 2) $y_{n+j} \equiv x_{m+j}$. In derselben Weise gibt es ein p und ein q, sodass 1) $(x_0 z_1 z_2 \cdots x_q)$ eine normale Kombination von $x_0 x_1 x_2 \cdots x_p$ ist, die weiterhin x_p wirklich enthält, und 2) $z_{q+j} \equiv x_{p+j}$. Wir können weiter annehmen, dass $p = n$ ist; denn ist $p > n$, so bleibt alles rechtig, das ich über m, n gesagt habe, wenn ich m durch $m + p - n$ ersetze, und ist $p < n$, so kann ich in ähnlicher Weise p durch n, auch q durch $q + n - p$ ersetzen.

Nach diesen Erklärungen sieht man sofort, dass u_i für $i \leqq q$ eine Kombination von $x_1, x_2, \cdots, x_m$ ist, während $u_{q+j} \equiv y_{n+j} \equiv x_{m+j}$. Daher ist $x_0 u_1 u_2 \cdots u_{q+1}$ eine normale Kombination von $x_1, x_2, \cdots, x_{m+1}$, und $x_0 u_1 u_2 u_3 \cdots$ ist die durch diese normale Kombination bestimmte Folge, w. z. b. w.

SATZ 2. *Sind Y und Z Kombinatoren, die den normalen Folgen η bzw. ζ von Variablen entsprechen, so entspricht $(Y \cdot Z)$ der Folge $(\eta \cdot \zeta)$.*

Beweis: Sind η und ζ, wie in der Festsetzung 4 bezeichnet, so gibt es m, n, p, q, sodass

$$Y x_0 x_1 x_2 \cdots x_m \doteq x_0 y_1 y_2 \cdots y_n \qquad x_m \text{ nicht ausgelassen}$$
$$Z x_0 x_1 x_2 \cdots x_p \doteq x_0 z_1 z_2 \cdots z_q \qquad x_p \text{ nicht ausgelassen.}$$

Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $n = p$ gilt; denn ist $p > n$, so können wir $x_{m+1}, x_{m+2}, \cdots, x_{m+p-n}$ zu den beiden Seiten der

ersten Gleichung hinzufügen, und ist $p < n$, so können wir $x_{p+1}, x_{p+2}, \cdots, x_n$ zu den beiden Seiten der zweiten Gleichung hinzufügen. Dann gilt

$$\begin{aligned}(Y \cdot Z)x_0x_1x_2 \cdots x_m &\doteq Y(Zx_0)x_1x_2 \cdots x_m \text{ (II B 4 Satz 1).} \\ &\doteq Zx_0y_1y_2y_3 \cdots y_n \\ &\doteq x_0u_1u_2 \cdots u_q,\end{aligned}$$

wo $u_i \equiv z_i$ mit x_i durch y_i ersetzt gilt.

§ 3. *Die Gruppierungen.*

Festsetzung 1. Eine Folge lauter Variablen heisst eine Gruppierung, wenn die Variablen darin in ihrer ursprünglichen Reihenfolge ohne Wiederholungen oder Auslassungen, aber natürlich in beliebiger Weise in Klammern zusammengefasst, erscheinen. Z. B. sind

$$x_0(x_1x_2)(x_3(x_4(x_5x_6)x_7))x_8x_9 \cdots$$
$$x_0(x_1(x_2(x_3x_4)x_5x_6x_7)x_8)x_9x_{10}$$

Gruppierungen. Jede Gruppierung ist eine normale Folge.

Festsetzung 2. Unter die Gruppierungen ist die Folge

$$x_0x_1x_2x_3 \cdots$$

einzuschliessen. Diese Gruppierung soll die *identische Gruppierung* heissen. Ihr entspricht der Identitätskombinator I.

Ich werde nun beweisen, dass jeder Gruppierung ein gewisser eindeutig bestimmter Kombinator entspricht.

SATZ 1. *Der Kombinator B_mB_n ($m \geqq 0$, $n > 0$) entspricht der Gruppierung, welche dann entsteht, wenn man $x_{m+1}, x_{m+2}, \cdots, x_{m+n+1}$ in einem einzigen Klammerpaar zusammenfasst. D. h.:*

$$B_mB_nx_0x_1x_2 \cdots x_{m+n+1} \doteq x_0x_1x_2 \cdots x_m(x_{m+1}x_{m+2} \cdots x_{m+n+1}).$$

Beweis:

$$\begin{aligned}B_mB_nx_0x_1x_2 \cdots x_{m+n+1} &\doteq B_n(x_0x_1x_2 \cdots x_m)x_{m+1} \cdots x_{m+n+1} \\ &\qquad \text{(vgl. II B 1, Satz 3),} \\ &\doteq x_0x_1x_2 \cdots x_m(x_{m+1}x_{m+2} \cdots x_{m+n+1}) \qquad \text{(vgl. II B 1, Satz 3).}\end{aligned}$$

SATZ 2. *Jeder Kombinator der Form*

$$(B_{m_q}B_{n_q}) \cdot (B_{m_{q-1}}B_{n_{q-1}}) \cdot (B_{m_{q-2}}B_{n_{q-2}}) \cdots (B_{m_2}B_{n_2}) \cdot (B_{m_1}B_{n_1}) \tag{1}$$

entspricht einer Gruppierung.

Beweis: Folgt aus Satz 1 und § 2 Satz 2, weil das Produkt (im Sinne von § 2) zweier Gruppierungen wieder eine Gruppierung ist.

SATZ 3. *Jeder Gruppierung, die nicht die identische ist, entspricht ein und nur ein Kombinator der Form* (1) *mit*

$$m_q > m_{q-1} > m_{q-2} > \cdots m_2 > m_1. \tag{2}$$

Beweis: Wir nehmen an, dass eine Gruppierung gegeben ist, worin alle die nach I C, Def. 1 fortgeschafften Klammern, sowie auch die die gesamte Gruppierung einschliessenden, wirklich fortgeschafft sind. Die übrig bleibenden Klammern befinden sich in Paaren—eine Anfangsklammer und eine ihr zugehörige Schlussklammer—ein solches Paar nennen wir ein Klammerpaar. Wir bezeichnen dann die Gruppierung mit Γ_q, wo q die Anzahl dieser übrig bleibenden Klammerpaare ist. Es gilt $q \geqq 1$, wenn die Gruppierung nicht die identische ist.

Nun sei das Klammerpaar, dessen Anfangsklammer am weitesten links steht, als das erste angesehen. Mit diesem verknüpfen wir die Zahlen m_1, n_1 wie folgt: x_{m_1} soll das letzte x sein, das vor der Anfangsklammer steht, und $n_1 + 1$ soll die Anzahl der innerhalb des Klammerpaares stehenden Glieder sein—wo ein eingeklammerter Teilausdruck, der selbst innerhalb eines anderen Klammerpaares steht, ist als ein einziges Glied des letzteren anzusehen.

Zunächst schaffen wir das erste Klammerpaar aus Γ_q fort. Die so gestaltete Gruppierung nennen wir Γ_{q-1}. Wir suchen dann das erste Klammerpaar in Γ_{q-1}, und bestimmen davon die Zahlen m_2 und n_2 genau so wie die vorigen m_1 und n_1 aus Γ_q bestimmt wurden. Dann schaffen wir dieses Klammerpaar weg und gestalten eine neue Gruppierung Γ_{q-2}, wovon wir die Zahlen m_3 und n_3 bestimmen, u. s. w.

Nachdem wir diesen Prozess q mal wiederholt haben, kommen wir auf einer Γ_0, welche keine Klammern enthält. Dann zeige ich, dass die so konstruierten Zahlen $m_1, m_2, \cdots m_q, n_1, n_2, \cdots n_q$ die Bedingungen des Satzes erfüllen.

Zunächst ist $m_{i+1} > m_i$. Nach der Definition ist $m_{i+1} \geqq m_i$, und die Gleichheit ist unmöglich, weil wir alle die nach I C Def. 1 erlaubten Klammerauslassungen ausgeführt haben, und also zwei Anfangsklammern an derselben Stelle nicht stehen können.

Zweitens: der Kombinator (1) mit diesem m_i und n_i entspricht dem Γ_q. In der Tat sei γ_r die Gruppierung, der $B_{m_r}B_{n_r}$ nach Satz 1 entspricht, dann folgt aus der Definition der Γ_i, dass

$$\begin{aligned} \Gamma_{r+1} &= \Gamma_r \cdot \gamma_{q-r} \qquad (r = 1, 2, \cdots, q-1) \\ \Gamma_1 &= \gamma_q \end{aligned}$$

gelten. Daher gilt (das Produkt von Folgen ist assoziativ)

$$\Gamma_q = \gamma_q \cdot \gamma_{q-1} \cdots \gamma_1.$$

Daraus folgt die Behauptung nach § 2, Satz 2.

Zuletzt gibt es nur einen Kombinator, der die Bedingungen erfüllt. Denn jeder andere Kombinator der Form (1), wofür (2) gilt, entspricht nach dem eben durchgeführten Beweis einer Gruppierung von ganz anderer Klammerstruktur. Aber derselbe Kombinator kann nicht zwei so verschiedenen Folgen entsprechen. (cf. § 1, Hilfsatz 3).

§ 4. *Die Umwandlungen.*

Festsetzung 1. Eine normale Folge von $x_0, x_1, x_2, \cdots, x_n$, worin nach den Auslassungen von I C Def. 1, keine Klammern (ausser den die gesamte Folge einschliessenden) erscheinen, nenne ich eine *Umwandlung*. (Diese Festsetzung stimmt mit der Erklärung im Abschnitte *A* überein). Z. B. sind

$$x_0 x_1 x_3 x_1 x_2 x_4 x_5 \cdots$$
$$x_0 x_2 x_4 x_2 x_3 x_5 x_6 \cdots$$

Umwandlungen, die erste ohne, die zweite mit Auslassungen.

Festsetzung 2. Die Folge:

$$x_0 x_1 x_2 \cdots$$

der der Kombinator I entspricht, ist sowohl eine Umwandlung als auch eine Gruppierung. Ich nenne sie die *identische Umwandlung*. Um weitere Umschreibungen zu vermeiden, soll hier festgestellt werden, dass diese identische Umwandlung zu allen den hierunter betrachteten Gattungen von Umwandlungen gehört.

SATZ 1. *Jede Umwandlung lässt sich in eindeutiger Weise als Produkt einer Umwandlung κ, die nur Auslassungen zulässt, wie etwa*

$$(1) \qquad (x_0 x_1 x_2 \cdots x_{h_1-1}\, x_{h_1+1} x_{h_1+2} \cdots x_{h_2-1} x_{h_2+1} \cdots x_{h_3-1} x_{h_3+1} \cdots x_{h_p-1} x_{h_p+1} \cdots),$$

und einer Umwandlung μ ohne Auslassungen darstellen.

Beweis: ω sei die gegebene Umwandlung. Wenn in ω keine Variablen ausgelassen werden, dann gilt $\omega = (\kappa \cdot \mu)$, wo κ die identische Umwandlung ist und $\mu = \omega$ ist. Sonst seien $x_{h_1}, x_{h_2}, \cdots, x_{h_p}$ die aus ω ausgelassenen Variablen. κ sei die Umwandlung (1), mit $h_1 \cdots h_p$ wie eben definiert. μ sei die Umwandlung, welche entsteht, wenn man in ω x_i durch x_j ersetzt, wo j aus i folgendermassen bestimmt wird: wenn $i < h_1$ ist, dann ist $j = i$; wenn $h_k < i < h_{k+1}$ ist, dann ist $j = i - k$; wenn $h_p < i$ ist, dann ist $j = i - p$. Dann ist μ eine Umwandlung ohne Auslassungen und $\omega = (\kappa \cdot \mu)$.

κ' sei nun irgendeine Umwandlung der Form (1) (bzw. die identische Umwandlung) und μ' sei eine Umwandlung ohne Auslassungen. Es sei $\omega' = (\kappa' \cdot \mu')$. Bilden wir κ'' und μ'' aus ω' genau wie wir κ und μ aus ω gebildet haben, so ist $\kappa'' = \kappa'$ und $\mu'' = \mu'$. Also wenn $\omega' = \omega$ gilt, so ist $\kappa' = \kappa$ und $\mu' = \mu$. Also sind κ und μ durch ω eindeutig bestimmt.

SATZ 2. *Jedem κ, das nicht das identische ist, entspricht ein und nur ein Kombinator der Form*

$$(2) \qquad K_{h_p} \cdot K_{h_{p-1}} \cdot \cdot \cdot K_{h_2} \cdot K_{h_1},$$

wo

$$(3) \qquad h_1 < h_2 < \cdot \cdot \cdot < h_{p-1} < h_p$$

gilt.

Beweis: Es sei eine Umwandlung κ der Form (1) gegeben. Der Kombinator (2) mit dem durch (1) bestimmten $h_1, h_2, \cdot \cdot \cdot, h_p$ entspricht dann diesem κ, und die Bedingung (3) ist natürlich erfüllt. Irgendein anderer Kombinator (2), wofür (3) erfüllt ist, entspricht nach dem eben Gesagten einer von κ verschiedenen Folge κ', also nicht zu κ (§ 1, Hilfssatz 3).

Festsetzung 3. Unter einer *Permutationsfolge* verstehen wir eine normale Folge, die durch eine Permutation bestimmt ist, oder, was dasselbe ist, eine Umwandlung ohne Auslassungen oder Wiederholungen.

SATZ 3. *Jede Umwandlung ohne Auslassungen lässt sich als Produkt zweier Faktoren darstellen, wovon der zweite eine Permutationsfolge ist, während im ersten die Variablen ihre ursprüngliche Reihenfolge behalten, aber wiederholt werden können. Dieser erste Faktor ist eindeutig bestimmt.*

Beweis: μ sei eine gegebene Umwandlung ohne Auslassungen. Wenn es in μ keine wiederholten Variablen gibt, so ist der zweite Faktor μ selbst, der erste die identische Umwandlung. Sonst seien $x_{k_1}, x_{k_2}, \cdot \cdot \cdot, x_{k_q}$ $(k_1 < k_2 < \cdot \cdot \cdot < k_q)$ sämtliche in μ wiederholte Variablen, und wir setzen fest, dass x_{k_1} $(r_1 + 1)$ mal, x_{k_2} $(r_2 + 1)$ mal u. s. w. bis x_{k_q} $(r_q + 1)$ mal in μ erscheinen. Dann betrachten wir die Umwandlung:

$$(4) \qquad (x_0 x_1 x_2 \cdots x_{k_1-1} x_{k_1} x_{k_1} \cdots (r_1 + 1)\,\mathrm{mal} \cdots x_{k_1} x_{k_1+1} \cdots x_{k_2-1} x_{k_2} x_{k_2} \cdots (r_2 + 1)\,\mathrm{mal} \\ \cdots x_{k_2} x_{k_2+1} \cdots x_{k_q-1} x_{k_q} x_{k_q} \cdots (r_q + 1)\,\mathrm{mal} \cdots x_{k_q} x_{k_q+1} \cdots).$$

Dann ist μ durch eine Permutation der in (4) erscheinenden Zeichen bestimmt, also ist es das Produkt von (4) und der durch diese Permutation bestimmten Permutationsfolge.

Umgekehrt sei eine Kombination (4) gegeben (wo wir unter $q = 0$ die

identische Umwandlung zu verstehen haben). Denn das Produkt von (4) nach irgendeiner Permutationsfolge ist ein μ, worin x_{k_i} $(i = 1, 2, \cdots, q)$ $(r_i + 1)$mal erscheint und kein anderes x wiederholt ist. Also kann ein μ nie zugleich als Produkt von zwei verschiedenen Kombinationen der Form (4) mit Permutationsfolgen dargestellt werden.

Der Satz wird nun bewiesen, wenn wir bemerken, dass wenn q, k_1, k_2, $\cdots, k_q$ beliebig sind, (4) die allgemeinste, den Bedingungen für den ersten Faktor genügende Folge ist.

Def. 1. $W_k{}^1 \equiv W_k \qquad k = 1, 2, 3, \cdots,$
$W_k{}^{r+1} \equiv W_k \cdot W_k{}^r \qquad k = 1, 2, 3, \cdots, \quad r = 1, 2, 3, 4, \cdots.$

SATZ 4. *Es gibt einen und nur einen Kombinator der Form*

(5) $$W_{k_q}{}^{r_q} \cdot W_{k_{q-1}}{}^{r_{q-1}} \cdots W_{k_2}{}^{r_2} \cdot W_{k_1}{}^{r_1},$$

wo ferner

(6) $$k_1 < k_2 < \cdots < k_q$$

*gilt, der einer gegebenen, von der identischen Umwandlung verschiedenen, den Bedingungen für den ersten Faktor im Satze 3 genügenden Folge entspricht.**

Beweis: Zuerst: $W_k{}^r$ entspricht der Folge

$$(x_0 x_1 x_2 \cdots x_{k-1} x_k x_k x_k \cdots (r+1)\text{mal} \cdots x_k x_{k+1} x_{k+2} \cdots).$$

In der Tat für $r = 1$ folgt dies aus II B 3, Satz 4. Ist es für ein gegebenes r angenommen, dann haben wir für $r + 1$

$$W_k{}^{r+1} x_0 x_1 x_2 \cdots x_k \doteq W_k{}^r x_0 x_1 \cdots x_{k-1} x_k x_k$$
$$\doteq x_0 x_1 x_2 \cdots x_{k-1} x_k x_k \cdots (r+2)\text{mal} \cdots x_k.$$

Es wird nun bewiesen werden, dass, wenn (6) gilt, (5) wie es geschrieben steht dem Ausdruck (4) entspricht. Zu diesem Behuf kürzen wir (5) mit q durch s ersetzt mit $\mathfrak{W}_s$, und den Ausdruck

$$(x_0 x_1 \cdots x_{k_1-1} x_{k_1} x_{k_1} \cdots (r_1 + 1)\text{mal}$$
$$\cdots x_{k_1} x_{k_1+1} \cdots x_{k_s-1} x_{k_s} x_{k_s} \cdots (r_s + 1)\text{mal} \cdots x_{k_s})$$

mit X_s ab. Dann haben wir schon für $s = 1$ bewiesen,

(7) $$\mathfrak{W}_s x_0 x_1 x_2 \cdots x_{k_s} \doteq X_s$$

Ist dies für ein bestimmtes s vorausgesetzt, so haben wir

* Dieser Satz und Lemma 3 meiner oben zit. Abhandlung sind wesentlich äquivalent. Der hier gegebene Beweis ist alternativ zu jenem.

$$\begin{aligned}
\mathfrak{W}_{s+1} x_0 x_1 x_2 \cdots x_{k_{s+1}} &\doteq W_{k_{s+1}}{}^{r_{s+1}} (\mathfrak{W}_s x_0) x_1 x_2 \cdots x_{k_{s+1}} \\
&\doteq \mathfrak{W}_s x_0 x_1 x_2 \cdots x_{k_{s+1}-1} x_{k_{s+1}} x_{k_{s+1}} \cdots (r_{s+1}) \text{mal} \cdots x_{k_{s+1}} \\
&\qquad \text{(nach dem eben bewiesenen),} \\
&\doteq X_s x_{k_s+1} x_{k_s+2} \cdots x_{k_{s+1}-1} x_{k_{s+1}} x_{k_{s+1}} \cdots (r_{s+1}+1) \text{mal} \cdots x_{k_{s+1}} \\
&\qquad \text{(nach der Voraussetzung),} \\
&\doteq X_{s+1} \qquad \text{(nach der Festsetzung über } X_s\text{).}
\end{aligned}$$

Also wird durch Induktion (7) für $s = q$, also die Behauptung bewiesen.

Der Beweis des Satzes folgt gleich. Denn wenn wir die Konstanten $q, k_1, k_2, \cdots, k_q$ in (5) einsetzen, so entspricht der resultierende Kombinator der Folge (4) nach dem letzten Absatz. Wenn wir andere Konstanten, die (6) genügen, in (5) einsetzen, so entspricht der resultierende Kombinator einer ganz anderen Folge der Form (4). Also gibt es nur einen Kombinator der betreffenden Beschaffenheit.

SATZ 5. *Jeder Permutationsfolge entspricht ein Kombinator $\mathfrak{C}$, der aus einem Produkt lauter $C_1, C_2, \cdots$ besteht, und zwar so, dass das mit dem höchsten Index versehene C_n nur einmal vorkommt.*

Beweis: Nach einem wohlbekannten Satz über Permutationen ist jede Permutation der Elemente $x_1, x_2, \cdots, x_m$ ein Produkt von Transformationen benachbarter Elementen. Dies bedeutet, in unsere Terminologie übersetzt, dass jede Permutationsfolge, die durch eine Permutation von $x_1, x_2, \cdots, x_m$ bestimmt wird, ein Produkt der Folgen

$$\begin{array}{l}
(x_0 x_2 x_1 x_3 \cdots) \\
(x_0 x_1 x_3 x_2 \cdots) \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
(x_0 x_1 x_2 \cdots x_{m-2} x_m x_{m-1} x_{m+1} \cdots)
\end{array}$$

ist. Diesen Folgen entsprechen bzw. die Kombinatoren $C_1, C_2, \cdots, C_{m-1}$. Infolgedessen entspricht der gegebenen Permutationsfolge ein $\mathfrak{C}$, das aus einem Produkt von lauter $C_1, C_2, \cdots, C_{m-1}$ besteht (§ 2, Satz 2).

Es sei nun eine Permutationsfolge π gegeben, die x_{m+1}, aber keine mit höherem Index versehene Variable, wirklich permutiert. x_k sei die Variable, die die Stelle von x_{m+1} einnimmt. Dann ist π ein Produkt von zwei Folgen π_1, und π_2, wo

$$\pi_1 = (x_0 x_1 x_2 \cdots x_{k-1} x_{k+1} \cdots x_{m-1} x_m x_{m+1} x_k x_{m+2} \cdots)$$

ist und π_2 durch eine Permutation von $x_1 x_2 \cdots x_m$ bestimmt ist. Der Folge π_1 entspricht aber der Kombinator $\mathfrak{C}_1$,

$$\mathfrak{C}_1 \equiv C_k \cdot C_{k+1} \cdots C_{m-1} \cdot C_m.$$

Der Folge π_2 entspricht weiter nach dem vorigen Absatz ein $\mathfrak{C}_2$, das ein Produkt lauter $C_1, C_2, \cdots C_{m-1}$ ist. Also entspricht $\mathfrak{C} \equiv \mathfrak{C}_1 \cdot \mathfrak{C}_2$ der Folge π, und $\mathfrak{C}$ erfüllt die Bedingungen des Satzes, weil C_m nur einmal vorkommt.

§ 5. *Darstellung der allgemeinen normalen Folge.*

SATZ 1. *Jede normale Folge lässt sich in eindeutiger Weise als Produkt einer Umwandlung und einer Gruppierung darstellen.*

Beweis: η sei die gegebene Folge. Wir erzeugen aus η eine Umwandlung ω und eine Gruppierung γ folgendermassen: zuerst schaffen wir alle die innerhalb η erscheinenden Klammern fort, dann soll der resultierende Ausdruck ω heissen. Zweitens lassen wir in η die Klammern stehen und schaffen die Variablen fort, und füllen dann die Leerstellen, wo Variablen früher waren, mit $x_0, x_1, x_2, \cdots$ von links nach rechts in ihrer naturgemässen Reihenfolge, ohne Auslassungen oder Wiederholungen aus. Der neue Ausdruck ist eine Gruppierung, γ. Diese ω und γ nennen wir die mit η assoziierte Umwandlung bzw. Gruppierung. Nach der Festsetzung 4, § 2 gilt $\eta \equiv (\omega \cdot \gamma)$.

Nun sei ω' irgendeine Umwandlung und γ eine Gruppierung. Es sei $\eta' \equiv (\omega' \cdot \gamma')$. Dann sind die mit η' assoziierte Gruppierung bzw. Umwandlung genau dieses ω' bzw. γ'. Infolgedessen muss, wenn $\eta' \equiv \eta$ ist, auch $\omega' \equiv \omega$ und $\gamma' \equiv \gamma$ sein.

SATZ 2. *Jeder normalen Folge entspricht mindestens ein Kombinator der Form:*

$$(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B}),$$

wo a) $\mathfrak{K}$ *in der Form von* § 4 *Satz* 2 *steht,*
b) $\mathfrak{W}$ *in der Form von* § 4 *Satz* 4 *steht,*
c) $\mathfrak{C}$ *in der Form von* § 4 *Satz* 5 *steht,*
d) $\mathfrak{B}$ *in der Form von* § 3 *Satz* 3 *steht.*

Ferner sind $\mathfrak{K}$, $\mathfrak{W}$ *und* $\mathfrak{B}$ *durch diese Bedingungen eindeutig bestimmt.*

Beweis: η sei eine gegebene normale Folge. Nach Satz 1 und § 4, Sätzen 1, 3 gibt es eine Gruppierung γ, eine Umwandlung κ, die nur Auslassungen zulässt, eine Umwandlung ω, die nur Wiederholungen zulässt, und eine Permutationsfolge, π, derart, dass

$$\eta \equiv (\kappa \cdot \omega \cdot \pi \cdot \gamma)$$

ist. Nach § 3 Satz 3, § 4 Sätze 2, 4, 5 gibt es gewisse die Bedingungen a-d erfüllende $\mathfrak{K}$, $\mathfrak{W}$, $\mathfrak{C}$, $\mathfrak{B}$ die diesen κ bzw. ω bzw. π bzw. γ entsprechen. Dann entspricht $(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B})$ der Folge η nach § 2 Satz 2.

Nun sei irgendeine der Folge η entsprechende Kombinator der Form (1) gegeben, etwa $(\mathfrak{K}' \cdot \mathfrak{W}' \cdot \mathfrak{C}' \cdot \mathfrak{B}')$. Es seien κ', ω', π', γ', die zu $\mathfrak{K}'$, $\mathfrak{W}'$, $\mathfrak{C}'$ bzw. $\mathfrak{B}'$ gehörenden Folgen; sie gehören denselben Kombinationsgattungen wie κ, ω, π bzw. γ an. Der betreffende Kombinator entspricht dann $(\kappa' \cdot \omega' \cdot \pi' \cdot \gamma')$ (§ 2, Satz 2); also, wenn er auch dem η entspricht, gilt

$$(\kappa' \cdot \omega' \cdot \pi' \cdot \gamma') = (\kappa \cdot \omega \cdot \pi \cdot \gamma) \qquad (\S\ 1 \text{ Hilfsatz } 3).$$

Also gelten $\gamma' = \gamma$ und $(\kappa' \cdot \omega' \cdot \pi) = (\kappa \cdot \omega \cdot \pi')$ (Satz 1); $\kappa' = \kappa$ und $(\omega' \cdot \pi') = (\omega \cdot \pi)$ (§ 4 Satz 1); $\omega' = \omega$ (§ 4 Satz 3). Daher sind $\mathfrak{B}'$ und $\mathfrak{B}$ identisch (§ 3 Satz 3); $\mathfrak{K}'$ und $\mathfrak{K}$ sind identisch (§ 2 Satz 2); $\mathfrak{W}'$ und $\mathfrak{W}$ sind identisch (§ 4 Satz 4). Also sind $\mathfrak{K}$, $\mathfrak{W}$, und $\mathfrak{B}$ durch die Bedingungen eindeutig bestimmt.

D. Reguläre Kombinatoren.

§ 1. *Vorläufige Festsetzungen und Sätze.*

Festsetzung 1. Ein Kombinator X heisst *regulär*, wenn er die Form

$$(X_1 \cdot X_2 \cdot \cdots \cdot X_n)$$

hat, wo jedes X_i ferner von einer der Formen

$$B_p B_q,\ C_q,\ W_q,\ K_q,\ B_p I,$$

ist. Die einzelne X_i heissen die *Glieder* von X.

Festsetzung 2. Ein Kombinator heisst *normal*, bzw. *in der normalen Form*, wenn er in der in II C 5 Satz 2 besprochenen Form steht.

Festsetzung 3. Ein Kombinator X heisst in einer gegebenen Form *umformbar*, wenn es ein schon in der betreffenden Form stehendes X' gibt, sodass $\vdash X = X'$.

In diesem Abschnitte beweise ich den Hauptsatz: wenn immer zwei reguläre Kombinatoren X und Y derselben Folge entsprechen, dann $\vdash X = Y$. Dies folgt daraus, dass erstens jeder reguläre Kombinator sich in die Normalform umformen lässt, und zweitens der Hauptsatz gilt, wenn nur X und Y normal sind. Der zweite in Abschnitt A erwähnte Hauptsatz wird hier für normale Kombinationen bewiesen.

Festsetzung 4. Zum Zwecke der Abkürzung möchte ich die folgenden Buchstaben für gewisse Gattungen regulärer Kombinatoren gebrauchen, derart, dass besondere Kombinatoren dadurch bezeichnet werden, dass Indizes an das betreffende Gattungszeichen angeheftet werden.

$\mathfrak{B}$: sämtliche Glieder der Form $B_m B_n$ oder $B_m I$.
$\mathfrak{C}$: sämtliche Glieder der Form C_n oder $B_m I$
$\mathfrak{K}$: sämtliche Glieder der Form K_n oder $B_m I$
$\mathfrak{M}$: sämtliche Glieder der Form W_n, C_n oder $B_m I$
$\mathfrak{W}$: sämtliche Glieder der Form W_n oder $B_m I$
Ω: sämtliche Glieder der Form C_n, K_n, W_n, oder $B_m I$.

Satz 1. *Zu jedem regulären Kombinator X gibt es ein X' derart, dass 1) X' regulär ist, 2) X' wenn von I selbst verschieden, gar keine Glieder der Form $B_m I$ enthält, während die anderen Glieder genau dieselben wie in X sind, und 3) $\vdash X = X'$.*

Beweis: Klar aus II B 2, Satz 1 und B 4, Satz 4. Wenn sämtliche Glieder in X der Form $B_m I$ sind, so ist X' gleich I, sonst ist X' von I verschieden.

Satz 2. $\vdash BB \cdot C_1 = C_1 \cdot C_2 \cdot B$.

Beweis: Wir haben zunächst aus Ax. $(CC)_1$ und II B 2, Satz 1

$$\vdash C_1 \cdot C_1 = B_2 I = I$$
$$\vdash C_2 \cdot C_2 = B(C_1 \cdot C_1) \qquad \text{(II B 4, Satz 2; II B 3, Def. 1).}$$
$$= B_3 I = I. \qquad (\text{Ax. } (CC)_1).$$

Daher:
$$\vdash BB \cdot C_1 = C_1 \cdot C_2 \cdot C_2 \cdot C_1 \cdot BB \cdot C_1$$
$$= C_1 \cdot C_2 \cdot B \cdot C_1 \cdot C_1 \qquad (\text{Ax. } (BC)),$$
$$= C_1 \cdot C_2 \cdot B. \qquad \text{w. z. b. w.}$$

§ 2. *Die kommutativen Gesetze.*

Festsetzung 1. Eine Gleichung der Form

$$\vdash C_1 B_{m+1} X = BX \cdot B_n$$

heisst ein *kommutatives Gesetz* für X, weil es eine gewisse Art von Vertauschbarkeit von X mit anderen Etwasen gewährt. Einige der Axiome sind von dieser Form; diese habe ich kommutative Axiome genannt.

Satz 1. *Wenn X ein Etwas ist, wofür*

$$\vdash CB_{m+1} X = BX \cdot B_n;$$

dann gilt für irgendein Etwas Y

$$\vdash B_m Y \cdot X = X \cdot B_n Y.$$

Beweis: Aus der Hp. und I D Satz 3 folgt

$$
\begin{array}{lll}
(1) & \vdash CB_{m+1}XY = (BX \cdot B_n)Y. & \\
\text{Aber} & \vdash CB_{m+1}XY = B_{m+1}YX & (\text{Reg. } C), \\
 & \quad = (B(B_mY))X & (\text{II B 1, Satz 5}), \\
(2) & \quad = B_mY \cdot X & (\text{II B 4, Def. 1}). \\
\text{Auch} & \vdash (BX \cdot B_n)Y = BX(B_nY) & (\text{II B 4, Satz 2}), \\
(3) & \quad = X \cdot B_nY. &
\end{array}
$$

aus (1), (2), (3) wird der Satz bewiesen.

SATZ 2. *Wenn X ein Etwas ist, wofür $\vdash CB_{m+1}X = BX \cdot B_n$; dann* $\vdash CB_{m+k+1}X = BX \cdot B_{n+k}$ $\quad (k = 1, 2, 3, \cdots)$.

Beweis: Wir haben zunächst mit Anwendungen der Eigenschaften der Gleichheit und Definitionen,

$$
\begin{array}{lll}
 & \vdash (CB_{m+1}X) \cdot B_k = (BX \cdot B_n) \cdot B_k & \\
(1) & \quad = BX \cdot B_{n+k} & (\text{II B 4, Sätze 3 und 5}).
\end{array}
$$

$$
\begin{array}{lll}
\text{aber} & \vdash CB_{m+1}X \cdot B_k = B(CB_{m+1}X)B_k & (\text{II B 4, Def. 1}), \\
 & \quad = BB(CB_{m+1})XB_k & (\text{Reg. } B), \\
(2) & \quad = (BB \cdot C)B_{m+1}XB_n & (\text{II B 4, Satz 1}), \\
 & \quad = (C_1 \cdot C_2 \cdot B)B_{m+1}XB_n & (\S\, 1, \text{Satz 2}), \\
 & \quad = C_1(C_2(BB_{m+1}))XB_n & (\text{II B 4, Sätze 1 und 3}), \\
 & \quad = C_2(BB_{m+1})\ B_nX & (\text{Reg. } C), \\
 & \quad = C_1(BB_{m+1}B_n)X & (\text{Def. von } C_2;\ \text{II B 3, Def. 1; Reg. } B) \\
(3) & \quad = CB_{m+k+1}X & (\text{II B 4, Def. 1 und Satz 5}).
\end{array}
$$

Aus (1) und (3) wird der Satz bewiesen.

SATZ 3. *Wenn X ein Etwas ist, wofür $\vdash CB_{m+1}X = BX \cdot B_n$; dann gilt für ein beliebiges Etwas Y*

$$(1) \qquad \vdash B_{m+k+h}Y \cdot B_hX = B_hX \cdot B_{n+k+h}Y \quad (k, h = 0, 1, 2 \cdots).$$

Beweis: Nach Satz 2 haben wir

$$\vdash CB_{m+k+1}X = BX \cdot B_{n+k}.$$

$\therefore$ nach Satz 1, $\vdash B_{m+k}Y \cdot X = X \cdot B_{n+k}Y$.

Die Behauptung (1) folgt dann aus II B 4, Satz 6 und II B 1, Satz 5.

SATZ 4. *Wenn X ein beliebiges Etwas ist; dann gilt für $m = 0, 1, 2, \cdots$, $n = 1, 2, \cdots$, und $p \geqq m + 1$,*

$$\vdash B_{p+n}Y \cdot B_mB_n = B_mB_n \cdot B_pY.$$

Beweis: Für $n = 1$ folgt der Satz aus Ax. B und Satz 3. Ist der Satz für ein gegebenes n angenommen, so wird er folgendermassen für $n + 1$ bewiesen:

$$\begin{aligned} \vdash B_{p+n+1}Y \cdot B_m B_{n+1} &= B_{p+n+1}Y \cdot B_m B_n \cdot B_m B && \text{(II B 4, Satz 7)},\\ &= B_m B_n \cdot B_{p+1}Y \cdot B_m B && \text{(Voraussetzung)},\\ &= B_m B_n \cdot B_m B \cdot B_p Y && \text{(dieser Satz für } n = 1),\\ &= B_m B_{n+1} \cdot B_p Y && \text{(II B 4, Satz 7)}. \end{aligned}$$

SATZ 5. *Wenn X ein beliebiges Etwas ist; dann sind die folgenden Gleichungen beweisbar,*

$$\begin{aligned} &\text{a)} \quad \vdash B_m Y \cdot C_p = C_p \cdot B_m Y, && \textit{wenn} \quad m \geqq p \geqq 1 \ \textit{gilt},\\ &\text{b)} \quad \vdash B_m Y \cdot W_p = W_p \cdot B_{m+1} Y, && \textit{wenn} \quad m \geqq p \geqq 1 \ \textit{gilt},\\ &\text{c)} \quad \vdash B_m Y \cdot K_p = K_p \cdot B_{m-1} Y, && \textit{wenn} \quad m \geqq p \geqq 1 \ \textit{gilt}. \end{aligned}$$

Beweis: Diese Gleichungen folgen aus Satz 3, den Axiomen C, W und K und den Definitionen von II B 3.

SATZ 6. *Das Axiom I_1 lässt sich aus den übrigen kombinatorischen Axiomen beweisen.*

Beweis:

$$\begin{aligned} \vdash CBI &= CB(WK) && \text{(I C, Def. 3)},\\ &= B(CB)WK && (\text{Reg. } B),\\ &= (B \cdot C)BWK && \text{(II B 4, Satz 1)},\\ &= (C_2 \cdot C_1 \cdot BB)BWK && (\text{Ax. } (BC)),\\ &= C_2(C_1(BBB))WK && \text{(II B 4, Satz 1)},\\ &= C_1(C_1 B_2 W)K && (\text{Def. von } C_2;\ \text{Reg. } B;\ \text{II B 1, Def. 1}),\\ &= C_1(B_2 W B_2)K && (\text{Ax. } W;\ \text{II B 4, Satz 1; II B 1, Satz 5}),\\ &= B_2 C_1 B_2 W B_2 K && \text{(II B 1, Satz 3)},\\ &= C_1 B_3 C_1 W B_2 K && (\text{Ax. } C;\ \text{II B 4, Satz 1; II B 1, Satz 5}),\\ &= B_3 W C_1 B_2 K && (\text{Reg. } C),\\ &= BW(C_1 B_2 K) && \text{(II B 1, Satz 2)},\\ &= BW(BK \cdot I) && (\text{Ax. } K),\\ &= W \cdot K_2 \cdot I && (\text{II B 4, Def. 1; Def. von } K_2),\\ &= W \cdot C_1 \cdot K \cdot I && (\text{Ax. } (CK)),\\ &= W \cdot K \cdot I && (\text{Ax. } (WC)),\\ &= BI \cdot I, \quad \text{w. z. b. w.} && (\text{Ax. } (WK)). \end{aligned}$$

§ 3. *Umformung in die Form $\Omega \cdot \mathfrak{B}$.*

SATZ 1. *Jedes $\mathfrak{B}$ kann entweder in I oder in die Normalform von II C 3, Satz 3, nämlich*

(1) $$B_{m_q}B_{n_q} \cdot B_{m_{q-1}}B_{n_{q-1}} \cdots B_{m_2}B_{n_2} \cdot B_{m_1}B_{n_1},$$
wo

(2) $$m_1 < m_2 < \cdots < m_q$$
gilt, umgeformt werden.

Beweis: Nach Festsetzung 4 und § 1, Satz 1 kann jedes $\mathfrak{B}$ entweder in I oder in die Form (1) umgeformt werden. Es bleibt nur zu beweisen, dass das betreffende $\mathfrak{B}$ im letzten Fall so umgeformt werden kann, dass auch (2) gilt. Ich beschränke mich auf solche $\mathfrak{B}$s.

Aus § 2, Satz 4 und II B 4, Satz 7 haben wir

(3) $$\vdash B_mB_n \cdot B_pB_q = B_{n+p}B_q \cdot B_mB_n, \quad \text{wenn} \quad p > m \text{ gilt.}$$

(4) $$\vdash B_mB_n \cdot B_pB_q = B_mB_{n+q}, \quad \text{wenn} \quad p = m \text{ gilt.}$$

Nun sei $\mathfrak{B}$ schon in der Normalform, dann kann $B_rB_s \cdot \mathfrak{B}$ in die Normalform umgeformt werden. In der Tat sei

$$r < m_t, m_{t+1}, \cdots, m_q \text{ und entweder } t = 1, \text{ oder } m_{t-1} \leqq r$$

Dann ist nach (3)

$$B_rB_s \cdot \mathfrak{B} = B_{m_q+s}B_{n_q} \cdot B_{m_{q-1}+s}B_{n_{q-1}} \cdots B_{m_t+s}B_{n_t} \cdot B_rB_s \cdot B_{m_{t-1}}B_{n_{t-1}} \cdots B_{m_1}B_{n_1},$$

wo natürlich, wenn $t = 1$ gilt, die Glieder rechts von B_rB_s an der rechten Seite nicht da sind. Wenn $t = 1$ oder $r > m_{t-1}$ gilt, ist die rechte Seite der eben Geschrieben schon in der Normalform. Sonst kann B_rB_s mit seinem rechtsstehenden Nachbarn nach (4) verschmolzen werden, und der neue Ausdruck wird in der Normalform sein.

Nun sei $\mathfrak{B}$ ein beliebiger Ausdruck der Form (1). $\mathfrak{B}_r$ sei $(r = 1, 2, \cdots)$ das Produkt der r rechtsstehenden Glieder von $\mathfrak{B}$. $\mathfrak{B}_1$ ist schon in der Normalform. Wenn $\mathfrak{B}_r$ in die Normalform umgeformt werden kann, so kann $\mathfrak{B}_{r+1} \equiv B_{m_{r+1}}B_{n_{r+1}} \cdot \mathfrak{B}_r$ nach dem vorigen Absatz in die Normalform umgeformt werden. Also kann $\mathfrak{B}_q \equiv \mathfrak{B}$ in die Normalform umgeformt werden.

SATZ 2. *Zu jedem $\mathfrak{B}$ und C_p gibt es ein $\mathfrak{B}'$ und ein $\mathfrak{C}'$ derart, dass*

$$\vdash \mathfrak{B} \cdot C_p = \mathfrak{C}' \cdot \mathfrak{B}'.$$

Beweis: Für $\mathfrak{B} \equiv I$, klar.
Zunächst sei $\mathfrak{B} \equiv B_mB$. Dann unterscheiden wir vier Fälle:

Fall 1: $p > m + 1$. Dann

$$\vdash B_mB \cdot C_p = C_{p+1} \cdot B_mB \qquad (\S\ 2, \text{ Satz } 4).$$

Fall 2: $p = m + 1$. Dann

$$\begin{aligned}
\vdash B_m B \cdot C_{m+1} &= B_m(B \cdot C_1) && \text{(II B 3, Satz 1; II B 4, Satz 6),}\\
&= B_m(C_2 \cdot C_1 \cdot BB) && \text{(Ax. }(BC)\text{),}\\
&= C_{m+2} \cdot C_{m+1} \cdot B_{m+1}B && \text{(II B 3, Satz 1; II B 1, Satz 5; II B 4, Satz 6).}
\end{aligned}$$

Fall 3: $p = m > 0$. Dann gilt

$$\begin{aligned}
\vdash B_m B \cdot C_m &= B_{m-1}(BB \cdot C)\\
&= B_{m-1}(C_1 \cdot C_2 \cdot B) && \text{(§ 1, Satz 2),}\\
&= C_m \cdot C_{m+1} \cdot B_{m+1}B && \text{(II B 3, Satz 1; II B 4, Satz 6).}
\end{aligned}$$

Fall 4: $p < m$. Dann gilt

$$\vdash B_m B \cdot C_p = C_p \cdot B_m B \qquad \text{(§ 2, Satz 5a).}$$

Also ist der Satz für $\mathfrak{B} \equiv B_m B$ bewiesen. Es folgt durch Induktion, dass es zu einem beliebigen $\mathfrak{C}$ ein $\mathfrak{C}'$ und ein $\mathfrak{B}'$ derart gibt, dass

$$\vdash B_m B \cdot \mathfrak{C} = \mathfrak{C}' \cdot \mathfrak{B}'.$$

Das allgemeinste $\mathfrak{B}$ kann nun entweder in I oder in ein Produkt von N Faktoren der Form $B_m B$ (Satz 1, II B 4, Satz 7)* umgeformt werden. Wir können nun ohne Beschränkung der Allgemeinheit $\mathfrak{B}$ als in dieser letzten Form gegeben betrachten. $\mathfrak{B}_M$ sei das Produkt der M rechtsstehenden Faktoren von $\mathfrak{B}$. Wenn der Satz für jedes $\mathfrak{B}_M$ bewiesen ist, so ist er für $\mathfrak{B}$ bewiesen. Aber für $M = 1$ ist er schon im letzten Absatz bewiesen. Für ein bestimmtes M sei angenommen, dass $\vdash \mathfrak{B}_M \cdot C_p = \mathfrak{C} \cdot \mathfrak{B}_M'$, dann gilt

$$\begin{aligned}
\vdash \mathfrak{B}_{M+1} \cdot C_p &= B_m B \cdot \mathfrak{B}_M \cdot C_p\\
&= B_m B \cdot \mathfrak{C} \cdot \mathfrak{B}_M'\\
&= \mathfrak{C}' \cdot \mathfrak{B}'' \cdot \mathfrak{B}_M'\\
&= \mathfrak{C}' \cdot \mathfrak{B}'.
\end{aligned}$$

Daher folgt der Satz durch Induktion für alle $\mathfrak{B}_M$, also auch für $\mathfrak{B}$.

SATZ 3. *Wenn X ein regulärer Kombinator ist, dessen sämtliche Glieder der form $B_m B_n$ oder C_p sind, so kann X in die Form* $(\mathfrak{C} \cdot \mathfrak{B})$ *umgeformt werden.*

Beweis: Es sei $X \equiv X_1 \cdot X_2 \cdots X_n$ wo die X_i die Glieder von X sind.

Es sei nun angenommen, dass $(X_1 \cdot X_2 \cdots X_q)$ in die Form $(\mathfrak{C}' \cdot \mathfrak{B}')$ umgeformt werden kann; dann gilt, wenn $X_{q+1} \equiv B_m B_n$,

* Für das $\mathfrak{B}$ von II C 3 (1) Satz 2 ist $N = n_1 + n_2 + n_3 + \cdots + n_q$.

$$\begin{aligned}\vdash X_1 \cdot X_2 \cdot X_3 \cdot \cdot \cdot X_q \cdot X_{q+1} &= \mathfrak{C}' \cdot \mathfrak{B}' \cdot B_m B_n \\ &= \mathfrak{C}' \cdot \mathfrak{B}'',\end{aligned}$$

während, wenn $X_{q+1} \equiv C_p$ ist, gilt

$$\begin{aligned}\vdash X_1 \cdot X_2 \cdot \cdot \cdot X_{q+1} &= \mathfrak{C}' \cdot \mathfrak{B}' \cdot C_p \\ &= \mathfrak{C}' \cdot \mathfrak{C}'' \cdot \mathfrak{B}'' && \text{(Satz 2)}, \\ &= \mathfrak{C}''' \cdot \mathfrak{B}''.\end{aligned}$$

Also ist der Satz durch Induktion auf q für X bewiesen, weil er für $q = 1$ klar ist.

SATZ 4. *Für jedes $\mathfrak{B}$ und W_p gibt es ein Ω und ein $\mathfrak{B}'$ derart, dass*

$$\vdash \mathfrak{B} \cdot W_p = \Omega \cdot \mathfrak{B}'.$$

Beweis: Für $\mathfrak{B} \equiv I$ klar. Ich beschränke mich also auf den Fall $\mathfrak{B} \not\equiv I$.

Zunächst zeige ich, dass es für jedes $m = 0, 1, 2, \cdot \cdot \cdot$, $p = 0, 1, 2, \cdot \cdot \cdot$, $k = 0, 1, 2, \cdot \cdot \cdot, p - 1$, ein $q > 0$, ein $h < q$ und ein X, deren sämtliche Glieder der Form $B_m B_n$ oder C_p sind, derart gibt, dass

$$\vdash B_m B \cdot W_p \cdot W_{p-1} \cdot \cdot \cdot W_{p-k} = W_q \cdot W_{q-1} \cdot \cdot \cdot W_{q-h} \cdot X. \tag{7}$$

Es sind drei Fälle zu unterscheiden:

Fall 1. $p \leqq m$. Weil nach § 2, Satz 5 für alle $r \leqq m$

$$\vdash B_m B \cdot W_r = W_r \cdot B_{m+1} B,$$

so haben wir hier,

$$\vdash B_m B \cdot W_p \cdot W_{p-1} \cdot \cdot \cdot W_{p-k} = W_p \cdot W_{p-1} \cdot \cdot \cdot W_{p-k} \cdot B_{m+k+1} B.$$

Fall 2. $p = m + 1$. Hier gilt für $k = 0$,

$$\begin{aligned}\vdash B_m B \cdot W_{m+1} &= B_m (B \cdot W) && \text{(II B 4, Satz 6)}, \\ &= B_m (W_2 \cdot W_1 \cdot C_2 \cdot B_2 B \cdot B) && \text{(Ax. } (BW)), \\ &= W_{m+2} \cdot W_{m+1} \cdot C_{m+2} \cdot B_{m+2} B \cdot B_m B && \text{(II B 4, Satz 6)}\end{aligned}$$

Für $k \geqq 1$,

$$\begin{aligned}\vdash B_m B \cdot W_{m+1} \cdot W_m \cdot \cdot \cdot W_{m-k+1} & \\ = W_{m+2} \cdot W_{m+1} \cdot C_{m+2} \cdot B_{m+2} B \cdot B_m B \cdot W_m \cdot \cdot \cdot W_{m-k+1} & \\ & \text{(nach dem Falle } k = 0), \\ = W_{m+2} \cdot W_{m+1} \cdot C_{m+2} \cdot W_m \cdot W_{m-1} \cdot \cdot \cdot W_{m-k+1} \cdot B_{m+k+2} B \cdot B_{m+k} B & \\ & \text{(nach Fall 1)}, \\ = W_{m+2} \cdot W_{m+1} \cdot W_m \cdot \cdot \cdot W_{m-k+1} \cdot C_{m+k+2} \cdot B_{m+k+2} B \cdot B_{m+k} B, &\end{aligned}$$

wobei der letzte Schritt aus § 2, Satz 5 und Definition von C_{m+k+2} folgt.

Fall 3. $p > m + 1$. Aus § 2, Satz 4 folgt, für $k = 0$,

$$\vdash B_mB \cdot W_p = W_{r+1} \cdot B_mB.$$

Also:

$$\begin{aligned}\vdash B_mB \cdot W_p \cdot W_{p-1} \cdots W_{m+3} \cdot W_{m+2} \cdot W_{m+1} \cdots W_{p-k},\\ = W_{p+1} \cdot W_p \cdots W_{m+3} \cdot W_{m+2} \cdot B_mB \cdot W_{m+1} \cdots W_{p-k},\\ = W_{p+1} \cdot W_p \cdots W_{m+3} \cdot W_{m+2} \cdots W_{p-k} \cdot C_{m+k+2} \cdot B_{m+k+2}B \cdot B_{m+k}B\end{aligned}$$

(Fall 2).

Also ist (7) bewiesen. Durch den Induktionsprozess, den ich im letzten Absatz des Beweises von Satz 2 benutzt habe, wird die Gleichung bewiesen, die entsteht wenn man in (7) B_mB durch ein beliebiges $\mathfrak{B}$ ersetzt. Wenn man in dieser Gleichung $k = 0$ setzt, so hat man

$$\vdash \mathfrak{B} \cdot W_p = W_q \cdot W_{q-1} \cdots W_{q-h} \cdot X,$$

wo, nach Satz 3 $\qquad \vdash X = \mathfrak{C} \cdot \mathfrak{B}'.$

Also $\qquad \vdash \mathfrak{B} \cdot W_p = W_q \cdot W_{q-1} \cdots W_{q-h} \cdot \mathfrak{C} \cdot \mathfrak{B}'$

$$= \Omega \cdot \mathfrak{B}' \qquad \text{w. z. b. w.,}$$

weil $(W_q \cdot W_{q-1} \cdots W_{q-h} \cdot \mathfrak{C})$ die Definition eines Ω erfüllt.

SATZ 5. *Für jedes $\mathfrak{B}$ und K_p gibt es ein $\mathfrak{K}$ und ein $\mathfrak{B}'$ derart, dass $\vdash \mathfrak{B} \cdot K_p = \mathfrak{K} \cdot \mathfrak{B}'$ oder auch, gibt es ein $\mathfrak{K}$, wofür $\vdash \mathfrak{B} \cdot K_p = \mathfrak{K}$.*

Beweis: Ich beweise den Satz zunächst für den Fall $\mathfrak{B} \equiv B_mB$. Dann gibt es drei Fälle:

Fall 1: $p \leqq m$. Dann gilt nach § 2, Satz 5

$$\vdash B_mB \cdot K_p = K_p \cdot B_{m-1}B.$$

Fall 2: $p = m + 1$. Dann folgt aus Ax. (BK) und II B 4, Satz 6

$$\vdash B_mB \cdot K_{m+1} = B_m(B \cdot K) = B_m(K_1 \cdot K_1) = K_{m+1} \cdot K_{m+1}.$$

Fall 3: $p > m + 1$. Dann folgt aus § 2, Satz 4,

$$\vdash B_mB \cdot K_p = K_{p+1} \cdot B_mB.$$

Der Rest des Beweises läuft genau wie in Satz 2.

SATZ 6. *Jeder reguläre Kombinator lässt sich in die Form $(\Omega \cdot \mathfrak{B})$ umformen.*

Beweis: X sei ein regulärer Kombinator und $X_1, X_2, \cdots, X_q$ seien seine Glieder, so dass

$$X \equiv X_1 \cdot X_2 \cdots X_n.$$

Der Satz ist sicher wahr für X_1. Nehmen wir an, er ist für den Kombinator $(X_1 \cdot X_2 \cdot \cdot \cdot X_q)$ wahr, dann werde ich ihn für $(X_1 \cdot X_2 \cdot \cdot \cdot X_{q+1})$ beweisen. In der Tat sei

$$\vdash X_1 \cdot X_2 \cdot \cdot \cdot X_q = \Omega' \cdot \mathfrak{B}'.$$

Dann ist X_{q+1} entweder $B_m B_n$, C_p, W_p oder K_p.* Im ersten Fall ist das zu Beweisende klar, wenn wir $\Omega \equiv \Omega'$, $\mathfrak{B} \equiv \mathfrak{B}' \cdot B_m B_n$ setzen. In anderen Fällen wissen wir aus den Sätzen 2, 4 und 5, dass es ein Ω'' und ein $\mathfrak{B}''$ gibt, wofür $\vdash \mathfrak{B}' \cdot X_{q+1} = \Omega'' \cdot \mathfrak{B}''$ gilt, also

$$\begin{aligned} \vdash X_1 \cdot X_2 \cdot \cdot \cdot X_{q+1} &= \Omega' \cdot \mathfrak{B}' \cdot X_{q+1} \\ &= \Omega' \cdot \Omega'' \cdot \mathfrak{B}'' \\ &= \Omega \cdot \mathfrak{B}, \end{aligned}$$

wenn wir $\Omega \equiv \Omega' \cdot \Omega''$, $\mathfrak{B} \equiv \mathfrak{B}''$ definieren.

§ 4. *Die Umformung* $\Omega = \mathfrak{K} \cdot \mathfrak{M}$.

SATZ 1. *Jedes* Ω *kann in die Form* $(\mathfrak{K} \cdot \mathfrak{M})$ *umgeformt werden.*

Beweis: Es genügt zu zeigen, dass jeder Kombinator der Form $(\mathfrak{M} \cdot K_p)$ in die betreffende Form übergeführt werden kann, denn das allgemeinste Ω enthält entweder kein K—und dann ist der Satz klar $(\mathfrak{K} \equiv I)$—, oder es kann in die Form

$$(\mathfrak{M}_1 \cdot K_{p_1} \cdot \mathfrak{M}_2 \cdot K_{p_2} \cdot \cdot \cdot \mathfrak{M}_k \cdot K_{p_k} \cdot \mathfrak{M}_{k+1})$$

umgeformt werden, wo einzelne $\mathfrak{M}_i \equiv I$ können. (In der Tat folgt dies durch Einschaltungen von gewissen I's, welche durch II B 4, Satz 4 erlaubt sind). Dann wird durch Wiederholung des Prozesses wodurch $(\mathfrak{M} \cdot K_p)$ in die Form des Satzes umgeformt wird, der ganze Ausdruck in diese Form gebracht.

Weiter genügt es zu beweisen, dass $W_m \cdot K_p$ in die Form $K_r \cdot W_s$ bzw. I und $C_m \cdot K_p$ in die Form $K_r \cdot C_s$ bzw. K_r umgeformt werden können. Denn wenn diese Behauptungen bewiesen sind, so folgt daraus, dass die einzelnen Glieder eines $\mathfrak{M}$ eins nach dem andern über die K's übertragen oder mit ihnen verschmolzen werden können.

Die Behandlung von $(C_m \cdot K_p)$ gibt vier Fälle:

Fall 1: $p \leqq m - 1$. Dann

$$\vdash C_m \cdot K_p = K_p \cdot C_{m-1} \qquad (\S\,2,\ \text{Satz 5c; II B 3}).$$

* Wir können natürlich annehmen, dass keine Glieder der Form $B_m I$ vorkommen, weil der Satz für $X \equiv I$ klar ist (s. § 1, Satz 1).

Fall 2: $p = m$. Dann

$$\begin{aligned} \vdash C_m \cdot K_m &= B_{m-1}(C_1 \cdot K_1) && \text{(II B 4, Satz 6; II B 3),} \\ &= B_{m-1}K_2 && \text{(Ax. } (CK)\text{),} \\ &= K_{m+1} && \text{(II B 3, Satz 5).} \end{aligned}$$

Fall 3: $p = m + 1$. Nach Fall 2 folgt

$$\begin{aligned} \vdash C_m \cdot K_{m+1} &= C_m \cdot C_m \cdot K_m \\ &= B_m(C_1 \cdot C_1) \cdot K_m && \text{(II B 3, Satz 1),} \\ &= B_{m+2}I \cdot K_m && \text{(Ax. } (CC)_1\text{; II B 1, Satz 5),} \\ &= K_m && \text{(II B 2, Satz 1, und II B 4, Satz 4).} \end{aligned}$$

Fall 4: $p > m + 1$. Dann nach § 2, Satz 5,

$$\vdash C_m \cdot K_p = K_p \cdot C_m.$$

Die Behandlung für $(W_m \cdot K_p)$ gibt drei Fälle:

Fall 1: $p \leqq m - 1$.

$$\vdash W_m \cdot K_p = K_p \cdot W_{m-1} \qquad \text{(§ 2, Satz 5c; II B 3, Sätze 3 u. 5).}$$

Fall 2: $p = m$.

$$\begin{aligned} \vdash W_m \cdot K_m &= B_{m-1}(W_1 \cdot K_1) && \text{(II B 4, Satz 6; II B 3, Sätze 3 u. 5),} \\ &= B_m I && \text{(Ax. } (WK)\text{; II B 1, Satz 5),} \\ &= I && \text{(II B 2, Satz 1).} \end{aligned}$$

Fall 3: $p > m$. Dann

$$\vdash W_m \cdot K_p = K_{p-1} \cdot W_m \qquad \text{(§ 2, Satz 5b; II B 3, Sätze 3 u. 5).}$$

Damit ist der Satz vollständig bewiesen.

SATZ 2. *Jedes $\mathfrak{K}$ kann entweder in I oder in die Normalform von II C 4, Satz 3, nämlich*

$$(K_{h_p} \cdot K_{h_{p-1}} \cdot \cdot \cdot K_{h_2} \cdot K_{h_1}) \tag{1}$$

wo

$$h_1 < h_2 < \cdot \cdot \cdot < h_p \tag{2}$$

sind, umgeformt werden.

Beweis: Nach § 1, Festsetzung 4 und § 1, Satz 1 kann $\mathfrak{K}$ entweder auf I oder auf die Form (1) gebracht werden. Aus § 2, Satz 5 folgt

$$\vdash K_m \cdot K_p = K_{p+1} \cdot K_m, \text{ wenn } p \geqq m. \tag{3}$$

Wenn es in dem betreffenden Ausdruck zwei benachbarte K's etwa K_{h_s} und $K_{h_{s-1}}$ gibt, wofür $h_{s-1} \geqq h_s$ ist, so kann eine gewisse Vertauschung stattfinden.

Nach einer gewissen Anzahl von Vertauschungen nach (3) wird der Ausdruck auf eine Form, wo (2) zutrifft, gebracht. Der genaue Beweis verläuft hier wie im § 3, Satz 1.

§ 5. *Die Normalform für* $\mathfrak{M}$.

In meiner oben erwähnten Abhandlung habe ich schon bewiesen, dass aus gewissen Axiomen (besser Axiomenschemen, wovon einige unendlich viele Axiome enthalten) die folgenden sich schliessen lassen: 1) jedes $\mathfrak{M}$ kann in die Normalform umgeformt werden, 2) wenn $\mathfrak{M}_1$ und $\mathfrak{M}_2$ derselben Folge entsprechen, so folgt $\vdash \mathfrak{M}_1 = \mathfrak{M}_2$. Um diese Ergebnisse unserer Theorie zu sichern, genügt es zu beweisen, dass die dort gegebenen Axiomen, und auch die Definitionen von W_2, $W_3 \cdots$ aus unserem Grundgerüst ableitbar sind.

Satz 1. $\vdash C_m \cdot C_m = I \qquad (m = 1, 2, 3 \cdots)$.

Beweis: Nach Definition von C_m und II B 4, Satz 6 gilt

$$\begin{aligned} \vdash C_m \cdot C_m &= B_{m-1}(C_1 \cdot C_1) & \\ &= B_{m-1}(B_2 I) & (\text{Ax. } (CC)_1), \\ &= I & (\text{II B 1, Satz 5; II B 2, Satz 1}). \end{aligned}$$

Satz 2. $\vdash C_m \cdot C_{m+1} \cdot C_m = C_{m+1} \cdot C_m \cdot C_{m+1} \qquad (m = 1, 2 \cdots)$.

Beweis: $$\begin{aligned} \vdash C_m \cdot C_{m+1} \cdot C_m &= B_{m-1}(C_1 \cdot C_2 \cdot C_1) & (\text{II B 4, Sätze 3 und 6}), \\ &= B_{m-1}(C_2 \cdot C_1 \cdot C_2) & (\text{Ax. } (CC)_2), \\ &= C_{m+1} \cdot C_m \cdot C_{m+1}. & \end{aligned}$$

Satz 3. $\vdash C_m \cdot C_{m+j} = C_{m+j} \cdot C_m$, *wenn* $j > 1$, $(m = 1, 2 \cdots)$.

Beweis: Folgt aus § 2, Satz 5, wenn wir C_{j-1} für Y in die Gleichung a) setzen.

Satz 4. $\vdash C_m \cdot W = W \cdot C_{m+1} \qquad (m = 2, 3, 4, \cdots)$.

Beweis: Folgt gleich aus § 2, Satz 5b.

Satz 5. $\vdash W_m \cdot W_n = W_n \cdot W_{m+1} \qquad (m \geqq n = 1, 2, 3, \cdots)$.

Beweis: Für $m = n$,

$$\begin{aligned} \vdash W_m \cdot W_m &= B_{m-1}(W_1 \cdot W_1) & (\text{II B 3, Def. 2; II B 4, Satz 6}), \\ &= B_{m-1}(W_1 \cdot W_2) & (\text{Ax. } (WW)), \\ &= W_m \cdot W_{m+1} & (\text{II B 3, Def. 2; II B 4, Satz 6}). \end{aligned}$$

Für $m > n$ folgt der Satz aus § 2, Satz 5b.

SATZ 6. $\vdash W_{m+1} = C_m \cdot W_m \cdot C_{m+1} \cdot C_m \qquad (m = 1, 2, 3 \cdots).$

Beweis: $\vdash C_m \cdot W_m = B_{m-1}(C_1 \cdot W_1)$ (II B 3, Def. 2; II B 4, Satz 6).
$= B_{m-1}(W_2 \cdot C_1 \cdot C_2)$ (Ax. (CW)),
$= W_{m+1} \cdot C_m \cdot C_{m+1}$ (II B 3, Def. 2; II B 4, Satz 6).

also $\vdash W_{m+1} = (W_{m+1} \cdot C_m \cdot C_{m+1}) \cdot C_{m+1} \cdot C_m$ (Satz 1),
$= C_m \cdot W_m \cdot C_{m+i} \cdot C_m$ w. z. b. w.

SATZ 7. *Wenn $\mathfrak{M}_1$ und $\mathfrak{M}_2$ derselben Folge lauter Variablen entsprechen, dann* $\vdash \mathfrak{M}_1 = \mathfrak{M}_2$.

Beweis: In meiner oben zitierten Abhandlung gegeben. Die Voraussetzungen jenes Beweises sind in der Tat schon hier bewiesen, wie folgt:

dort		hier
Axiomschema	I	Satz 1
"	II	Satz 2
"	III	Satz 3
"	IV	Satz 4
"	V	Ax. (WC)
"	VI	Folgen aus Satz 5 durch Umkehrung des Beweises der
"	VII	Gleichungen (6) und (7) meiner zitierten Abhandlung.
Definition von W_k		Satz 6.

Jener Beweis lässt sich aber vermöge der hier vorliegenden Entwicklungen bedeutend abkürzen. In der Tat können wir aus § 2, Satz 5b und Ax. (CW) in einer den Beweisen von § 3, Sätzen 2 und 4 ähnlicher Weise schliessen, dass ein $\mathfrak{M}$ in die Form $\mathfrak{W} \cdot \mathfrak{C}$ umgeformt werden kann, und dann weiter, wie im § 3, Satz 1 nachweisen, dass $\mathfrak{W}$ sich in die Normalform umformen lässt. Dabei werden Lemmas 1 und 2 jener Abhandlung bewiesen. Für Lemmas 3 und 4 sind alternative Beweise schon in II C 4, Sätzen 4 und 5 geliefert.

SATZ 8. *Jedes $\mathfrak{M}$ lässt sich in die Normalform umformen.*

Beweis: Dies ist im Laufe des Beweises von Satz 7 dargetan. (Lemmas 1 und 2 meiner früheren Abhandlung).—Der Satz folgt auch direkt aus Satz 7, § 6 (unten) Satz 2, und II C 5, Satz 2.

§ 6. *Zusammenfassung und Schluss.*

SATZ 1. *Jeder reguläre Kombinator kann in die Normalform umgeformt werden.*

Beweis: Jeder reguläre Kombinator X lässt sich in die Form $(\Omega \cdot \mathfrak{B})$, wo $\mathfrak{B}$ in der Normalform steht, umformen (§ 3, Sätze 1 und 6). Dieses Ω lässt sich in die Form $(\mathfrak{K} \cdot \mathfrak{M})$ umformen, wo $\mathfrak{K}$ in der Normalform ist (§ 4, Sätze 1 und 2). Endlich lässt sich $\mathfrak{M}$ in die Normalform umformen (§ 4, Satz 8). Also kann X in die Normalform $(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B})$ umgeformt werden.

SATZ 2. *Jeder reguläre Kombinator entspricht einer normalen Folge lauter Variablen, und zwar im ersten Sinne.*

Beweis: Die einzelnen Glieder eines regulären Kombinator entsprechen solchen Folgen (II B 3; II C 3, Satz 1; II B 2). Daher entspricht das Produkt einer solchen Folge (II C 2, Satz 2). Dass er der Folge im ersten Sinne entspricht, ist aus dem Beweis von II C 2, Satz 2 ohne weiteres ersichtlich.

SATZ 3. *Wenn X_1 und X_2 reguläre Kombinatoren sind, wofür* $\vdash X_1 = X_2$; *dann sind X_1 und X_2 äquivalent in dem dritten Sinne.*

Beweis: Nach II C 1, Satz 11, und Satz 2 sind X_1 und X_2 im ersten Sinne äquivalent, also entsprechen sie beide einer gemeinsamen Folge lauter Variablen. Nach Satz 2 entsprechen sie dieser Folge im ersten Sinne. Also ist der Sinn der Äquivalenz zwischen X_1 und X_2 der dritte.

SATZ 4. *Wenn X_1 und X_2 reguläre, derselben Folge von lauter Variablen entsprechende Kombinatoren sind; dann* $\vdash X_1 = X_2$.

Beweis: Sind Y_1 und Y_2 reguläre, in der Normalform stehende Kombinatoren, in welche X_1 bzw. X_2 umgeformt werden können (Satz 1), so entsprechen Y_1 und Y_2 derselben Folge wie X_1 und X_2 (Hp. und Satz 3).

Sind	$Y_1 \equiv \mathfrak{K}_1 \cdot \mathfrak{M}_1 \cdot \mathfrak{B}_1$ und $Y_2 \equiv \mathfrak{K}_2 \cdot \mathfrak{M}_2 \cdot \mathfrak{B}_2$,	
dann	$\vdash \mathfrak{B}_1 = \mathfrak{B}$ und $\vdash \mathfrak{K}_1 = \mathfrak{K}_2$	(II C 5, Satz 2),
und	$\vdash \mathfrak{M}_1 = \mathfrak{M}_2$	(§ 5, Satz 7).
Daher	$\vdash Y_1 = Y_2$.	
also	$\vdash X_1 = X_2$,	w. z. b. w.

SATZ 5. *Damit für zwei reguläre Kombinatoren X_1 und X_2* $\vdash X_1 = X_2$ *gilt, ist es notwendig und hinreichend, dass X_1 und X_2 im dritten Sinne äquivalent sind.*

Beweis: Klar aus Sätzen 3 und 4.

Festsetzung 1. Eine normale Folge ξ hat die Ordnung n, wenn 1) es

eine Kombination X von $x_0, x_1, x_2, \cdots, x_n$ gibt, sodass die Folge durch (Xx_{n+1}) bestimmt ist, und 2) n die keinste Zahl ist, wofür ein solches X existiert.

Es folgt aus dieser Festsetzung, dass jeder Kombinator der dem ξ entspricht, ihm mindestens mit der Ordnung $n+1$ entspricht.*

SATZ 6. *ξ sei eine normale Folge der Ordnung n, und X sei ein der Folge ξ entsprechender normaler Kombinator. Dann entspricht X der Folge ξ mit der Ordnung $n+1$.*

Beweis: Wir nehmen ein m so gross, dass X', wobei

$$X' \equiv Xx_0x_1x_2 \cdots x_m,$$

sich auf einen Abschnitt von ξ reduziert. Wenn in dieser Reduktion keine der Variablen $x_{n+1}, x_{n+2}, \cdots, x_m$ gestört werden, so ist der Satz bewiesen. Sonst führen wir die Reduktion von X' ohne Störung von $x_{n+1}, x_{n+2}, \cdots, x_m$ soweit fort, bis wir auf einen Ausdruck der Form

$$Y(Zx_0)y_1y_2 \cdots y_q$$

kommen (wo Y ein Glied von X_k, Z † ein Produkt solcher Glieder ist, und $y_1 \cdots y_q$ Kombinationen von $x_1x_2 \cdots x_m$ sind), sodass eine weitere Reduktion auf einen Ausdruck derselben Form ohne Störung von $x_{n+1} \cdots x_m$ nicht möglich ist. Wir unterscheiden dann vier Fälle:

1) Y ist ein K_p. Dann wird ein x_s, $s > n$, in der weiteren Reduktion ausgelassen. Weil durch Reduktionsprozesse keine Variablen eingesetzt werden, so bleibt x_s ausgelassen bis zur Ende der Reduktion von X'. Weil dieses x_s nicht in ξ ausgelassen ist, kann X nicht der Folge ξ entsprechen.

2) Y ist ein W_p. Dann wird in der weiteren Reduktion ein x_s, $s > n$, verdoppelt. Weil X normal ist, so kann kein Glied der Form K_p in Z vorkommen; also bleibt x_s verdoppelt bis zur Ende. Weil x_s nicht in ξ verdoppelt ist, so kann X auch in diesem Falle nicht der Folge ξ entsprechen.

3) Y ist ein C_p. In diesem Falle führen wir die Reduktion fort, bis wir an einen Ausdruck der obigen Form ankommen, wo nun Y das C_p mit höchstem Index ist. Durch dieses C_p wird ein höchstes x_s $s > n$ mit einer niedrigeren x_t vertauscht, und weil dieses C_p nur einmal vorkommt (§ 1, Fest-

* Wir haben hier $n+1$, nicht n, weil ich die Variable x_0 zugelassen habe. Die Behauptung folgt, weil in jeder Reduktion auf einen Abschnitt von ξ die Variable x_n gestört werden muss.

† Streng genommen, können wir statt (Zx_0) einen Ausdruck haben, worauf (Zx_0) sich reduziert; aber dies stört den Kern des Beweises gar nicht.

setzung 2), so kann x_s nie seine Stelle wieder erreichen. Aber dies widerspricht noch einmal der Voraussetzung, dass X der Folge ξ entspricht.

4) Y ist ein $B_p B_q$. Dann reduziert sich X' auf eine Kombination, worin mindestens ein x_s, $s > n$, eingeklammert ist. Daher entspricht X nicht der Folge ξ.

Diese vier Fälle erschöpfen alle möglichkeiten, weil Glieder der Form $B_m I$ in einem normalen Kombinator nicht vorkommen.

Satz 7. *Wenn X eine beliebige normale Kombination von lauter Variablen ist, so gibt es einen normalen Kombinator, der sie darstellt.*

Beweis: Wir nehmen an, dass X eine normale Kombination der Variablen $x_0, x_1, \cdots, x_n$ ist. Y sei der normale Kombinator, welcher der durch X bestimmten Folge entspricht (II C 5, Satz 2). Die Ordnung dieses Entsprechens ist $\leqq n+1$ (Satz 6, Festsetzung 1). Also muss $(Y x_0 x_1 x_2 \cdots x_n)$ sich aus X reduzieren, und daher wird ipso facto X durch Y dargestellt.

E. Eigentliche Kombinatoren.

§ 1. *Vorläufige Festsetzungen und Sätze.*

Festsetzung 1. Ein Kombinator heisst *eigentlich,* wenn er einer Folge lauter Variablen entspricht.

In diesem Abschnitte beweise ich, dass jeder eigentliche Kombinator in der Form $\mathfrak{R}I$, wo $\mathfrak{R}$ regulär ist, umgeformt werden kann. Daraus folgt, hinsichtlich der Ergebnisse des letzten Abschnitts, dass zwei derselben Folge entsprechende Kombinatoren immer gleich sind. Der Beweis der in Abschnitt A erwähnten Hauptsätze II und III wird hier vollzogen (der letzte für eigentliche Kombinatoren).

Festsetzung 2. Ausser den Gattungszeichen von II D 1, Festsetzung 4 benutze ich den Buchstaben $\mathfrak{R}$ für einen regulären Kombinator.

Festsetzung 3. Ein Kombinator heist *regulierbar,* wenn er in einen regulären Kombinator umgeformt werden kann; d. h. wenn es einen regulären Kombinator gibt, der ihm gleich ist.

Satz 1. *Sind die Kombinatoren X und Y regulierbar, so ist auch $(X \cdot Y)$ regulierbar.*

Beweis: Nach den Voraussetzungen gibt es $\mathfrak{R}_1$ und $\mathfrak{R}_2$, sodass $\vdash X = \mathfrak{R}_1$ und $\vdash Y = \mathfrak{R}_2$, also $\vdash X \cdot Y = \mathfrak{R}_1 \cdot \mathfrak{R}_2$. $(\mathfrak{R}_1 \cdot \mathfrak{R}_2)$ ist aber regulär (dies folgt direkt aus II D 1, Festsetzung 1).

SATZ 2. *Ist der Kombinator X regulierbar, so ist jedes $(\mathfrak{B}X)$ regulierbar.*

Beweis: Wenn $\mathfrak{B} \equiv I$ ist, klar.

Zunächst sei $\mathfrak{B} \equiv B_m$. Setzen wir dann

$$\vdash X = \mathfrak{N}, \quad \mathfrak{N} \equiv X_1 \cdot X_2 \cdot \cdot \cdot X_n.$$

Dann $\qquad \vdash B_m X = B_m X_1 \cdot B_m X_2 \cdot \cdot \cdot B_m X_n \qquad$ (II B 4, Sätze 3 u. 6),

und die rechte Seite ist regulär.

Es sei nun ein allgemeines $\mathfrak{B}$ gegeben. Wir können annehmen, dass $\mathfrak{B}$ in der Normalform steht. Dann folgt wenn $m_1 = 0$ ist (wo m_1 wie in II C 3, Satz 3 zu verstehen ist),

$$\vdash \mathfrak{B} = B\mathfrak{B}' \cdot B_{n_1}$$

also
$$\vdash \mathfrak{B}X = B\mathfrak{B}'(B_{n_1}X) = \mathfrak{B}' \cdot B_{n_1}X.$$

Die rechte Seite ist regulierbar nach dem eben Bewiesenen und Satz 1 Dagegen sei $m_1 > 0$. Dann

$$\vdash \mathfrak{B}X = B_{m_1}\mathfrak{B}'X = B(B_{m_1-1}\mathfrak{B}')X$$
$$= B_{m_1-1}\mathfrak{B}' \cdot X.$$

Die rechte Seite ist wieder regulierbar nach dem oben Gesagten und Satz 1.

SATZ 3. *Wenn X und Y beliebige Etwase sind, dann $\vdash XY = (X \cdot BY)I$.*

Beweis: Klar aus II B 2, Satz 4, und II B 4, Satz 1.

SATZ 4. *Jeder Kombinator der Form $(\mathfrak{N}I)$ entspricht einer Folge lauter Variablen, und zwar in dem ersten Sinne.*

Beweis: n sei so gewählt, dass der Ausdruck $(\mathfrak{N}x_0x_1x_2 \cdot \cdot \cdot x_n)$ sich auf eine normale Kombination von $x_0, x_1, x_2, \cdot \cdot \cdot, x_n$, etwa $(x_0y_1y_2 \cdot \cdot \cdot y_q)$ ohne Auslassung von x_n reduziert (möglich nach II D 6, Satz 2). Dann wird $(\mathfrak{N}Ix_1x_2 \cdot \cdot \cdot x_n)$ auf $(Iy_1y_2 \cdot \cdot \cdot y_q)$ im ersten Sinne reduziert. Dass sich die weitere Reduktion auf $(y_1y_2 \cdot \cdot \cdot y_q)$ im ersten Sinne vollzieht, ist selbstverständlich. Also entspricht $(\mathfrak{N}I)$ der durch die eben geschilderte Kombination bestimmte Folge.

SATZ 5. *Eine notwendige und hinreichende Bedingung dafür, dass ein $(\mathfrak{N}I)$ einer normalen Folge entspricht, ist, dass es ein $\mathfrak{N}'$ und ein $\mathfrak{B}$ gibt, sodass*

$$\vdash \mathfrak{N} = B\mathfrak{N}' \cdot \mathfrak{B}.$$

Beweis: Die Bedingung ist hinreichend; denn ist sie erfüllt, so gilt

$$\vdash \mathfrak{N}I = (B\mathfrak{N}' \cdot \mathfrak{B})I = \mathfrak{N}' \cdot \mathfrak{B}I.$$

($\mathfrak{B}I$) ist regulierbar nach Satz 2; also ist ($\mathfrak{R}I$) regulierbar nach Satz 1. Daher entspricht ($\mathfrak{R}I$) einer normalen Folge (Satz 4; II D 6, Satz 2; II C 1, Satz 11).

Die Bedingung ist notwendig. In der Tat sei angenommen, dass $(\mathfrak{R}Ix_1x_2 \cdots x_n)$ sich auf eine normale Kombination V von $x_1, x_2, \cdots, x_n$ reduziert. Dann erscheint x_1 in V vereinzelt und an der ersten Stelle. $\mathfrak{R}$ werde in die Normalform umgeformt, etwa

$$\vdash \mathfrak{R} = \mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B}.$$

Dann ist $\mathfrak{K}$ von der Faktor K_1 frei, weil sonst x_1 in V ausfallen würde, also

$$\vdash \mathfrak{K} = B\mathfrak{K}'.$$

Gleichfalls ist $\mathfrak{W}$ von der Faktor W_1 frei, weil sonst x_1 in V verdoppelt sein würde, also $\vdash \mathfrak{W} = B\mathfrak{W}'$. Weiter entspricht $\mathfrak{C}$ einer durch eine Permutation der Variablen $x_2, x_3, \cdots, x_m$ bestimmten Folge, also ist $\mathfrak{C}$ in ein Produkt von $C_2, C_3, \cdots, C_{n-1}$, umformbar* und daher $\vdash \mathfrak{C} = B\mathfrak{C}'$. Aus den letzten drei Formeln folgt

$$\vdash \mathfrak{R} = B(\mathfrak{K}' \cdot \mathfrak{W}' \cdot \mathfrak{C}') \cdot \mathfrak{B} \qquad \text{w. z. b. w.}$$

SATZ 6. *Zu jeder Folge lauter Variablen gibt es ein $\mathfrak{R}_1$, und zwar ein normales $\mathfrak{R}_1$ ohne Glieder der Form B_n, sodass ($\mathfrak{R}_1I$) der Folge entspricht. Gibt es überdies ein anderes der Folge entsprechendes $\mathfrak{R}_2$, so gilt für ein durch $\mathfrak{R}_2$ bestimmtes n*

$$\vdash \mathfrak{R}_2 = \mathfrak{R}_1 \cdot B_n.$$

Beweis: Wir nehmen an, die Variablen in der gegebenen Folge sind $x_1, x_2, x_3, \cdots$. Die Folge sei etwa

(1) $$x_jy_1y_2y_3 \cdots \qquad j \geqq 1.$$

wo y_i eine Kombination gewissen x's ist. $\mathfrak{R}_1$ sei ein normaler Kombinator, welcher der Folge

(2) $$x_0x_jy_1y_2y_3 \cdots$$

entspricht. Dann entspricht ($\mathfrak{R}_1I$) der gegebenen Folge nach dem Beweis von Satz 4. Enthält $\mathfrak{R}_1$ ein Glied der Form B_n, so müsste $\mathfrak{R}_1$, weil es normal ist, von der Form $(\mathfrak{R}_1' \cdot B_n)$ sein; aber in diesem Falle würde $\mathfrak{R}_1$ einer Folge entsprechen, worin eine Anfangsklammer links von der zweiten Variablen steht. Weil (2) diese Form nicht hat, so erfüllt $\mathfrak{R}_1$ die Bedingungen des ersten Teils des Satzes.

---

* Vgl. Beweis von II C 4, Satz 5.

Nun sei $\mathfrak{R}_2$ irgendein regulärer Kombinator derart, dass $(\mathfrak{R}_2 I)$ der gegebenen Folge entspricht. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $\mathfrak{R}_2$ normal ist (II D 6, Satz 1). Wenn $\mathfrak{R}_2$ Glieder der Form B_n enthält, so gibt es ein $\mathfrak{R}_2'$ ohne solche Glieder, und ein B_n, sodass

$$\vdash \mathfrak{R}_2 = \mathfrak{R}_2' \cdot B_n.$$

Dann gilt $\vdash \mathfrak{R}_2 I = \mathfrak{R}_2'(B_n I) = \mathfrak{R}_2' I$ (II B 2, Satz 1; II B 2, Satz 1). Im entgegengesetzten Falle setzen wir $\mathfrak{R}_2' \equiv \mathfrak{R}_2$. $\mathfrak{R}_2'$ entspricht in den beiden Fällen einer Folge der Form

$$x_0 x_k z_1 z_2 z_3 \cdots,$$

(d. h. ohne Klammern vor der zweiten Variable.) Daher entspricht $\mathfrak{R}_2 I$ nach dem Beweis von Satz 4 der Folge:

$$x_k z_1 z_2 z_3 \cdots.$$

Weil dies mit der gegebenen Folge übereinstimmen muss, so ist $k = j$, $z_1 = y_1$, $z_2 = y_2$ u. s. w. $\mathfrak{R}_2'$ entspricht daher derselben Folge wie $\mathfrak{R}_1$. Also:

$$\vdash \mathfrak{R}_2' = \mathfrak{R}_1 \qquad \text{(II D 6, Satz 4).}$$
$$\therefore \quad \vdash \mathfrak{R}_2 = \mathfrak{R}_1 \cdot B_n \qquad \text{w. z. b. w.}$$

Festsetzung 4. Eine von der Variablen x_0 frei Folge ξ heisst der Ordnung n, wenn 1) es eine Kombination X von $x_1, x_2, \cdots x_n$ gibt, sodass die Folge durch $X x_{n+1}$ bestimmt wird, 2) n die kleinste Zahl dieser Beschaffenheit ist.

SATZ 7. *Dass $(\mathfrak{R}_1 I)$ von Satz 6 entspricht seiner Folge mit einer Ordnung, die mit der Ordnung der Folge selbst übereinstimmt.*

Beweis: Das $\mathfrak{R}_1$ entspricht seiner normalen Folge mit der Ordnung $n+1$, wo n die Ordnung der Folge selbst ist. (II D 6, Satz 6). Wie im Satz 4 folgt daraus, dass $(\mathfrak{R}_1 I)$ seiner Folge mit der Ordnung n entspricht.

SATZ 8. *Zu jeder Kombination lauter Variablen gibt es mindestens einen Kombinator, der sie darstellt.*

Beweis: Folgt aus Sätzen 6 und 7.

§ 2. *Die Kombinatoren* Γ *und eine Verallgemeinerung der kommutativen Gesetze.* Diese Sätze sind Hilfssätze für § 3 unten.

Def. 1. $\Gamma_1 \equiv C_1; \quad \Gamma_{n+1} \equiv \Gamma_n \cdot C_{n+1}, \qquad (n = 1, 2, 3, \cdots).$

SATZ 1. $\vdash \Gamma_n = C_1 \cdot C_2 \cdots C_n.$

Beweis: Klar.

SATZ 2. *Wenn $X_0, X_1, \cdots, X_n, Y$ beliebige Etwase sind, so gilt* $\vdash \Gamma_n X_0 Y X_1 X_2 \cdots X_n = X_0 X_1 \cdots X_n Y$.

Beweis: Für $n = 1$, klar aus Regel C.

Ist nun der Satz für ein bestimmtes n angenommen, dann wird er für $n + 1$ wie folgt bewiesen:

$$
\begin{aligned}
\vdash \Gamma_{n+1} X_0 Y X_1 \cdots X_{n+1} &= \Gamma_n (C_{n+1} X_0) X_1 Y X_2 \cdots X_n && \text{(Def. 1; II B 4, Satz 1)},\\
&= C_{n+1} X_0 X_1 \cdots X_n Y X_{n+1} && \text{(Voraussetzung)},\\
&= X_0 X_1 \cdots X_n X_{n+1} Y && \text{(II B 3, Satz 2)}.
\end{aligned}
$$

Also folgt der Satz durch Induktion.

SATZ 3. $\vdash \Gamma_{n+1} = C_1 \cdot B\Gamma_n$.

Beweis: Für $n = 1$ klar.

Ist der Satz für ein bestimmtes n angenommen, so gilt für dieses n

$$
\begin{aligned}
\vdash \Gamma_{n+2} &= \Gamma_{n+1} \cdot C_{n+2} && \text{(Def. 1)},\\
&= C_1 \cdot B\Gamma_n \cdot C_{n+2} && \text{(Hp.)},\\
&= C_1 \cdot B(\Gamma_n \cdot C_{n+1}) && \text{(II B 3, Def. 1; II B 4, Sätze 2, 3)},\\
&= C_1 \cdot B\Gamma_{n+1} && \text{(Def. 1)}.
\end{aligned}
$$

Also folgt der Satz durch Induktion.

SATZ 4. $\vdash BB \cdot \Gamma_n = \Gamma_{n+1} \cdot B$.

Beweis: Für $n = 1$ ist dies in II D 1, Satz 2 bewiesen.

Ist der Satz für ein bestimmtes n angenommen, dann

$$
\begin{aligned}
\vdash BB \cdot \Gamma_{n+1} &= BB \cdot \Gamma_n \cdot C_{n+1} && \text{(Def. 1)},\\
&= \Gamma_{n+1} \cdot B \cdot C_{n+1} && \text{(Hp.)},\\
&= \Gamma_{n+1} \cdot C_{n+2} \cdot B && \text{(II D 2, Satz 4)},\\
&= \Gamma_{n+2} \cdot B && \text{(Def. 1)}.
\end{aligned}
$$

Also wird der Satz durch Induktion bewiesen.

SATZ 5. *Wenn Y ein beliebiges Etwas ist; dann*

$$\vdash B_p(C_1 B_{n+1} Y) = \Gamma_{p+1}(B_p B_{n+1}) Y, \qquad (p = 0, 1, 2, 3 \cdots).$$

Beweis: Definieren wir vorübergehend

$$X_p \equiv \Gamma_p (B_{p-1} B_{n+1}) Y, \qquad (p = 1, 2, 3 \cdots);$$

dann $\quad X_1 \equiv C_1 B_{n+1} Y,$

und

$$\begin{aligned}
\vdash BX_p &= B_2B\Gamma_p(B_{p-1}B_{n+1})Y && \text{(II B 1, Satz 3)},\\
&= (BB \cdot \Gamma_p)(B_{p-1}B_{n+1})Y && \text{(II B 1, Satz 5; II B 4, Def. 1)},\\
&= \Gamma_{p+1}(B(B_{p-1}B_{n+1}))Y && \text{(Satz 4; II B 4, Satz 1)},\\
&= \Gamma_{p+1}(B_pB_{n+1})X && \text{(II B 1, Satz 5)},\\
&= X_{p+1}.
\end{aligned}$$

Also folgt der Satz aus II B 1, Satz 4.

SATZ 6. *Wenn X, Y beliebige Etwase sind, so gilt*

$$\vdash \Gamma_{p+1}(B_pB_{n+1} \cdot X)Y = \Gamma_{p+2}(B_{p+1}B_{n+1})YX.$$

Beweis:

$$\begin{aligned}
\vdash \Gamma_{p+1}(B_pB_{n+1} \cdot X)Y &= \Gamma_{p+1}(B_{p+1}B_{n+1}X)Y && \text{(II B 4, Def. 1; II B 1, Satz 5)},\\
&= B\Gamma_{p+1}(B_{p+1}B_{n+1})XY && \text{(Reg. } B),\\
&= (C_1 \cdot B\Gamma_{p+1})(B_{p+1}B_{n+1})YX && \text{(II B 4, Satz 1; Reg. } C),\\
&= \Gamma_{p+2}(B_{p+1}B_{n+1})YX. && \text{(Satz 3)}.
\end{aligned}$$

SATZ 7. *Wenn XY Etwase sind, und Y das Kommutativgesetz*

$$\vdash C_1B_{m+1}Y = BY \cdot B_n$$

erfüllt; dann

$$\vdash \Gamma_{p+1}(B_pB_{m+1} \cdot X)Y = B_{p+1}Y \cdot B_pB_n \cdot X.$$

Beweis: Nach den Voraussetzungen,

$$\begin{aligned}
\vdash \Gamma_{p+1}(B_pB_{m+1} \cdot X)Y &= \Gamma_{p+2}(B_{p+1}B_{m+1})YX && \text{(Satz 6)},\\
&= B_{p+1}(C_1B_{m+1}Y)X && \text{(Satz 5)},\\
&= B_p(C_1B_{m+1}Y) \cdot X && \text{(II B 1, Satz 5; II B 4, Def. 1)},\\
&= B_p(BY \cdot B_n) \cdot X && \text{(Hp.)},\\
&= B_{p+1}Y \cdot B_pB_n \cdot X && \text{(II B 4, Satz 6; II B 1, Satz 5)}.
\end{aligned}$$

§ 3. *Darstellung der allgemeinen Kombinationen.*

Festsetzung 1. Ein Ausdruck X der Form

$$(\mathfrak{R}IY_1Y_2 \cdots Y_px_1x_2 \cdots x_n),$$

wo die Y_i Etwase sind, *reduziert sich formal* auf einen Ausdruck Z, wenn mit Behandlung der Y_i als Variablen eine Reduktion von X auf Z sich durchführen lässt; oder, falls man es genauer haben will, wenn der Ausdruck $(\mathfrak{R}Ix_1x_2 \cdots x_{n+p})$ sich auf ein solches Z' reduziert, dass durch Einsetzung von Y_i statt x_i für $i = 1, 2, \cdots, p$, und von x_{i-p} statt x_i für $i = p+1, p+2, \cdots, p+n$ in Z', der Ausdruck Z erzielt wird.

SATZ 1. *Ist X ein Kombinator, so gibt es ein S der Form ($\mathfrak{R}IBCWK$), das sich auf X formal reduziert.*

Beweis: Ersetzen wir in dem gegebenen Kombinator B, C, W, K durch x_1, x_2, x_3 bzw. x_4, so erzeugen wir eine Kombination Z von x_1, x_2, x_3, x_4. Nach § 1, Sätzen 6 und 7 gibt es ein $\mathfrak{R}$, sodass ($\mathfrak{R}Ix_1x_2x_3x_4$) sich auf Z reduziert. Daher reduziert ($\mathfrak{R}IBCWK$) sich formal auf X, w. z. b. w.

SATZ 2. *X sei eine Kombination von Kombinatoren und Variablen $x_1, x_2, \cdots, x_n$, sodass*

a) *X auf einen ähnlichen X' durch einen einzigen Reduktionsprozess reduziert wird,*

b) *es ein S gibt, näm.*

$$S \equiv \mathfrak{R}IY_1Y_2 \cdots Y_p, \tag{A}$$

wo jedes Y_i entweder B, C, W oder K ist, sodass der Ausdruck ($Sx_1x_2 \cdots x_n$) sich auf X formal reduziert.

Dann gibt es ein S', nämlich

$$S' \equiv \mathfrak{R}'IY_1'Y_2' \cdots Y_q' \tag{B}$$

wo jedes Y_i' entweder B, C, W oder K ist, sodass

α) $\vdash S' = S$,

β) *der Ausdruck ($S'x_1x_2 \cdots x_n$) sich auf X' formal reduziert.*

Beweis: $Y_0', Y_1', Y_2', \cdots, Y_q'$ seien die sämtlichen in X vorkommenden Grundkombinatoren (B, C, W oder K), und zwar so, dass jeder der Kombinatoren B, C, W, K unter diesen Y_i' genau so oft erscheint, wie in X selbst. Die Anordnung dieser Kombinatoren unter den Y_i' bleibt für jetzt gleichgültig. Die Y_i' kommen natürlich—abgesehen von ihrer Häufigkeit—unter den $Y_1, Y_2, \cdots, Y_p$ vor.

Behandeln wir nunmehr die $Y_1, Y_2, \cdots, Y_p$ formal als Variable, so schliessen wir die Folgenden:

1) Die Folge $IY_0'Y_1' \cdots Y_q'x_1x_2x_3 \cdots$ ist eine Umwandlung der Folge $IY_1Y_2 \cdots Y_px_1x_2x_3 \cdots$,

2) Wenn wir die durch X bestimmte Folge wie folgt schreiben, $y_0y_1y_2y_3 \ldots$, wo y_0 entweder ein Y_i oder eine Variable ist, und y_i, $i > 0$, eine Kombination von $Y_1, Y_2, \cdots Y_p$ und Variablen ist, so ist die Folge

$$Iy_0y_1y_2 \cdots \tag{1}$$

das Produkt der eben erwähnten Umwandlung und eine Folge derselben Form wie (1).

Nun bezeichne ich mit Ω bzw. $\mathfrak{R}_1$ zwei normale Kombinatoren, sodass Ω bzw. $\mathfrak{R}_1 I$ dieser Umwandlung bzw. der zuletzt erwähnten Folge entsprechen. Dann bemerken wir: 1) $(\Omega \cdot \mathfrak{R}_1)$ entspricht der Folge (1) (II C 2, Satz 2); 2) wir dürfen annehmen dass $\mathfrak{R}_1$ und also $(\Omega \cdot \mathfrak{R}_1)$* kein Glied der Form B_n enthält (weil $\mathfrak{R}_1$ normal ist und $(\mathfrak{R}_1 I)$ einer Folge der Form (1) entspricht—vgl. Beweis von § 1, Satz 6); 3) wir dürfen ferner annehmen, dass $\mathfrak{R}$ kein Glied der Form B_n hat (denn wenn $\vdash \mathfrak{R} = \mathfrak{R}^* \cdot B_n$, $\mathfrak{R}^*$ normal, so können wir in den Satz $\mathfrak{R}$ durch $\mathfrak{R}^*$ ersetzen). Daraus folgt

$$\begin{aligned} &\vdash \mathfrak{R} = \Omega \cdot \mathfrak{R}_1 && (\S\ 1, \text{ Satz } 6). \\ \therefore\quad &\vdash S = (\Omega \cdot \mathfrak{R}_1) I Y_1 Y_2 \cdots Y_p && \\ &\quad = \mathfrak{R}_1 I Y_0' Y_1' \cdots Y_q' && (\text{nach der Bedeutung von } \Omega). \end{aligned}$$

Weiterhin reduziert der Ausdruck

$$(\mathfrak{R}_1 I Y_0' Y_1' \cdots Y_q' x_1 x_2 \cdots x_n) \tag{2}$$

sich formal auf X. (Nach der Bedeutung von $\mathfrak{R}_1$, § 1, Satz 7).

Wir unterscheiden nun zwei Fälle; näm.—

I. Die Reduktion von X auf X' volzieht sich in dem ersten Sinne.

II. Die Reduktion von X auf X' vollzieht sich in dem zweiten Sinne.

Fall I. Hier sei Y_0' der erste in X vorkommende Grundkombinator. Dann erscheint Y_0' in X nur an der ersten Stelle. Deshalb ist X eine normale Kombination von $Y_0', Y_1', Y_2', \cdots, Y_q'$ und Variablen. Es gibt also, nach § 1, Satz 5, ein $\mathfrak{R}_2$ und ein $\mathfrak{B}$, sodass

$$\begin{aligned} &\vdash \mathfrak{R}_1 = B\mathfrak{R}_2 \cdot \mathfrak{B}. \\ \therefore\quad &\vdash \mathfrak{R}_1 I Y_0' = (\mathfrak{R}_2 \cdot \mathfrak{B} I) Y_0' \\ &\quad = (\mathfrak{R}_2 \cdot \mathfrak{B} I \cdot B Y_0') I \qquad (\S\ 1, \text{ Satz } 3). \end{aligned} \tag{3}$$

Nun betrachten wir Y_0' wieder als Kombinator und definieren:

a) $\mathfrak{R}' \equiv$ eine normale Form von $(\mathfrak{R}_2 \cdot \mathfrak{B} I \cdot B Y_0')$† ohne Glieder dor Form B_n,

b) $$S' \equiv \mathfrak{R}' I Y_1' Y_2' \cdots Y_q',$$

so folgt $\vdash S' = S$. Also ist die Bedingung α) erfüllt.

Dieses $\mathfrak{R}' I$ entspricht, wenn wir $Y_1', Y_2', \cdots, Y_q'$ formal betrachten, der durch X' bestimmten Folge. Denn ich habe gezeigt, dass der Ausdruck (2)

* Sogar wenn es auf die Normalform gebracht wird.

† Dies ist regulär nach § 1, Sätzen 1 und 2.

sich formal auf X reduziert. In dieser Reduktion betrachten wir Y_0' nunmehr nicht als Variable, sondern als Kombinator; dabei wird nichts in der Reduktion geändert. Die Reduktion lässt sich doch eine Stufe weiter auf X' durchführen (nach Hp. a). Aber weil

$$\vdash \mathfrak{R}' = \mathfrak{R}_1 I Y_0' \qquad \text{(aus (3))},$$

und die beiden Seiten dieser Gleichung Folgen lauter Variablen entsprechen, so entsprechen sie derselben Folge (II C 1, Satz 11).

Dass die Bedingung β) erfüllt ist, folgt daraus nach § 1, Satz 7.

Fall II. Hier soll Y_0' den Kombinator bezeichnen, welcher durch die Reduktion von X auf X' eliminiert wird. Er nehme in X die $(r+1)$te Stelle ein, wo $r > 0$ nach der Voraussetzung dieses Falles ist.

Nach II D 6, Satz 1, gibt es $\mathfrak{K}_1$, $\mathfrak{W}_1$, $\mathfrak{C}_1$, und $\mathfrak{B}_1$ derart, dass

$$(4) \qquad \vdash \mathfrak{R}_1 = \mathfrak{K}_1 \cdot \mathfrak{W}_1 \cdot \mathfrak{C}_1 \cdot \mathfrak{B}_1.$$

Aber nach der Voraussetzung über $Y_0', Y_1', \cdots, Y_q'$ kann $\mathfrak{K}_1$ kein K_i für $i \leqq q+1$ und $\mathfrak{W}_1$ kein W_j für $j \leqq q+1$ enthalten, also wird

$$(5) \qquad \vdash \mathfrak{K}_1 \cdot \mathfrak{W}_1 = B_{q+1}(\mathfrak{K}_2 \cdot \mathfrak{W}_2).$$

Auch entspricht $\mathfrak{C}_1$ einer Permutationsfolge, welche in zwei Faktoren zerlegt werden kann, wie folgt: der erste Faktor lässt Y_0' invariant, aber ordnet $Y_1', Y_2', \cdots, Y_q'$ und die Variablen in die Anordnung, die sie in X haben, an; der zweite Faktor setzt Y_0' an die Stelle, die es in X hatt, aber lässt die Anordnung von $Y_1', \cdots, Y_q'$ und die Variablen unter sich selbst, unverändert bleiben. Dem ersten Faktor entspricht ein $\mathfrak{C}$, dessen Glieder alle C_i mit $i > 1$ sind, also ein $\mathfrak{C}$ von der Form $B\mathfrak{C}_2$; dem zweiten Faktor entspricht Γ_r $(r > 0)$. Also (II D 5, Satz 7).

$$(6) \qquad \vdash \mathfrak{C}_1 = B\mathfrak{C}_2 \cdot \Gamma_r.$$

Daher (aus (4) (5) (6))

$$(7) \qquad \vdash \mathfrak{R}_1 = B(B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W}_2 \cdot \mathfrak{C}_2) \cdot \Gamma_r \cdot \mathfrak{B}.$$

Nun erscheint Y_0' nach Hp. (a) und Definition am Anfang eines in X eingeklammerten Teilausdrucks; die Anzahl der Glieder ausser Y_0' dieses Teilausdrucks sei $m+1$. Dann (vgl. den Beweis von II C 3, Satz 3, und II D 3, Satz 1) gibt es $\mathfrak{B}_2$ und $\mathfrak{B}_3$ derart, dass

$$(8) \qquad \vdash \mathfrak{B} = B_{r+1}\mathfrak{B}_2 \cdot B_r B_{m+1} \cdot \mathfrak{B}_3.$$

Weil die Glieder von Γ_r alle $C_1, C_2, \cdots$ oder C_r sind, so kann $B_{r+1}\mathfrak{B}_2$ mit allen diesen Gliedern, also mit Γ_r selbst, vertauscht werden (II D 2, Satz 5a).

Daher $\vdash \mathfrak{N}_1 = B(B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W} \cdot \mathfrak{C}_2) \cdot \Gamma_r \cdot B_{r+1}\mathfrak{B}_2 \cdot B_r B_{m+1} \cdot \mathfrak{B}_3,$

$$= B(B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W}_2 \cdot \mathfrak{C}_2 \cdot B_r\mathfrak{B}_2) \cdot \Gamma_r \cdot B_r B_{m+1} \cdot \mathfrak{B}_3$$

(9) $$= B\mathfrak{N}_2 \cdot \Gamma_r \cdot B_r B_{m+1} \cdot \mathfrak{B}_3.$$

wenn ich nur definiere:

$$\mathfrak{N}_2 \equiv B_q\mathfrak{K}_2 \cdot B_q\mathfrak{W}_2 \cdot \mathfrak{C}_2 \cdot B_r\mathfrak{B}_2.$$

Daher $\vdash \mathfrak{N}_1 I Y_0' = (B\mathfrak{N}_2 \cdot \Gamma_r \cdot B_r B_{m+1} \cdot \mathfrak{B}_3) I Y_0'$

$$= B\mathfrak{N}_2(\Gamma_r((B_r B_{m+1} \cdot \mathfrak{B}_3) I)) Y_0'$$

(10) $$= \mathfrak{N}_2(\Gamma_r(B_{r-1} B_{m+1} \cdot \mathfrak{B}_3 I) Y_0').$$

Weil nach Hp. a) und Definition von Y_0' eine Reduktion durch Y_0' wirklich stattfindet, so muss $m \geqq 2$ sein, wenn $Y_0'B$ oder C ist, und $m \geqq 1$, wenn $Y_0'W$ oder K ist. Infolgedessen muss es nach II D 2, Satz 2, und den kommutativen Axiomen ein n geben, wofür $\vdash C_1 B_{m+1} Y_0' = BY_0' \cdot B_n$. Also

$$\vdash \Gamma_r(B_{r-1} B_{m+1} \cdot \mathfrak{B}_3 I) Y_0' = B_r Y_0' \cdot B_{r-1} B_n \cdot (\mathfrak{B}_3 I) \qquad (\S\, 2,\ \text{Satz } 7);$$

also, wenn wir dies in (10) einsetzen,

$$\vdash \mathfrak{N}_1 I Y_0' = \mathfrak{N}_2(B_r Y_0' \cdot B_{r-1} B_n \cdot \mathfrak{B}_3 I)$$

(11) $$= (\mathfrak{N}_2 \cdot B_{r+1} Y_0' \cdot B_r B_n \cdot \mathfrak{B}_3) I \qquad (\S\, 1,\ \text{Satz } 3).$$

Definiere ich nun

a) $$\mathfrak{N}' \equiv \mathfrak{N}_2 \cdot B_{r+1} Y_0' \cdot B_r B_n \cdot \mathfrak{B}_3,$$

b) $$S' \equiv \mathfrak{N}' I Y_1' Y_2' \cdots Y_q',$$

so folgt aus (11) und (A), dass $\vdash S = S'$.

Dass die Bedingung β) erfüllt ist, folgt hier genau wie im Fall I.

SATZ 3. *Ist X ein solcher Kombinator, dass*

(1) $$(X x_1 x_2 \cdots x_n)$$

sich auf eine Kombination von $x_1, x_2, \cdots, x_n$ reduziert; dann lässt X sich in eine $(\mathfrak{N} I)$ umformen und zwar so, dass $(\mathfrak{N} I x_1 x_2 \cdots x_n)$ sich auf die gegebene Kombination reduziert.

Beweis: Nach den Voraussetzungen gibt es eine Reihe von Ausdrücken $X_1, X_2, \cdots, X_m$ derart, dass 1) X_{i+1} sich aus X_i durch einen einzigen Reduktionsprozess erzielt, 2) X_1 mit dem Ausdruck (1) identisch ist, 3) X_m eine Kombination von $x_1, x_2, \cdots, x_n$ ist.

Wir können nun diesen X_i eine Reihe von Kombinatoren $S_1, S_2, \cdots, S_m$ zuordnen und zwar so dass

a) Jedes S_i in der Form (A) (s. Satz 2) steht,
b) $(S_i x_1 x_2 \cdots x_n)$ sich auf X_i formal reduziert,
c) $\vdash S_{i+1} = S_i$.

In der Tat gilt als S_1 der in Satz 1 ausgestellte Kombinator; und aus Satz 2 folgt, dass aus einem gegebenen S_i, $(i < m)$ ein S_{i+1} konstruiert werden kann.

In dieser Weise haben wir ein S_m, etwa

$$(2) \qquad S_m \equiv \mathfrak{R}_m I Y_1 Y_2 \cdots Y_p \quad (Y_i \equiv B, C, W \text{ oder } K, \mathfrak{R}_m \text{ normal})$$

sodass $(S_m x_1 x_2 \cdots x_n)$ sich *formal* auf eine Kombination lauter Variablen reduziert. In dieser Reduktion müssen freilich alle $Y_1, Y_2, \cdots, Y_p$ ausfallen. Also wenn $\mathfrak{R}_m$ auf die normale Form gebracht wird, gilt

$$\vdash \mathfrak{R}_m = K_p \cdot K_{p-1} \cdots K_1 \cdot \mathfrak{R}_m'.$$

Infolgedessen $\quad \vdash S_m = \mathfrak{R}_m' I \quad$ (aus (2), II B 3).

Aber $\quad \vdash S_m = S_1 \quad$ (aus c)),

$\quad = X \quad$ (Bedeutung von S_1).

$\therefore \quad \vdash X = \mathfrak{R}_m' I. \quad$ w. z. b. w.

SATZ 4. *Wenn zwei Kombinatoren Y_1 und Y_2 derselben Folge lauter Variablen entsprechen;*

dann $\quad \vdash Y_1 = Y_2.$

Beweis: Nach Satz 3 gibt es $\mathfrak{R}_1$ und $\mathfrak{R}_2$, sodass

$$\vdash Y_1 = \mathfrak{R}_1 I \qquad \vdash Y_2 = \mathfrak{R}_2 I,$$

und die beiden Kombinatoren $(\mathfrak{R}_1 I)$ und $(\mathfrak{R}_2 I)$ auch derselben Folge entsprechen. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $\mathfrak{R}_1$ und $\mathfrak{R}_2$ normal und ohne Glieder der Form B_n sind.

Dann $\quad \vdash \mathfrak{R}_1 = \mathfrak{R}_2 \quad$ (§ 1, Satz 6).

Also $\quad \vdash Y_1 = Y_2. \quad$ w. z. b. w.

SATZ 5. *Wenn zwei eigentliche Kombinatoren Y_1 und Y_2 dieselbe Kombination von lauter Variablen darstellen, dann $\vdash Y_1 = Y_2$.*

Beweis: Klar aus Satz 4.

§ 4. *Die Substitutionsprozesse.*

Zum Schluss gebe ich hier einige Sätze über die Verhältnisse der Substitutionsprozesse zu den Kombinatoren. Die Bewiese gebe ich nur kurz, weil sie meistens nur Rechnungsübungen sind.

Die Substitutionsprozesse lassen sich zunächst durch Kombinationen von Variablen darstellen. Z. B. betrachten wir den Ausdruck:

$$(u x_1 (v x_2 x_3) x_4).$$

Wenn u und v Grundfunktionen sind, so bedeutet dies eine gewisse aus einer

Verknüpfung von u und v erzeugte Funktion von x_1, x_2, x_3, x_4. Aber wir können ihn auch,—wenn wir u und v für Variablen halten—als eine Funktion von u und v betrachten, welche für bestimmte Werte von u und v jene Funktion von x_1, x_2, x_3, x_4 darstellt—d. h. als den Verknüpfungsprozess selbst. Diese Auffassung ist naturgemäss, weil nach der Ausdeutung von Anwendung der Ausdruck für irgendeine bestimmten Werte von u, v, x_1, x_2, x_3, x_4 die mit der Auffassung verträgliche Aussage bedeutet. Der Ausdruck lässt sich ferner in

$$(C_1 \cdot BB_2)uvx_1x_2x_3x_4$$

umformen. Von unserem Gesichtspunkte aus ist also $(C_1 \cdot BB_2)$ der Substitutionsprozess selbst—eine Funktion, welche aus u und v die Funktion $((C_1 \cdot BB_2)uv)$ liefert, wo diese letzte die Funktion ist, welche aus x_1, x_2, x_3 die eben geschilderte Aussage liefert. In diesem Sinne können wir sagen, dass $(C_1 \cdot BB_2)$ den betreffenden Substitutionsprozess *darstellt.*

Von diesem Gesichtspunkt aus haben wir die folgenden Satze:

SATZ 1. *Jede Umwandlung im Sinne von Abschnitt A lässt sich durch ein Ω darstellen.*

SATZ 2. *Die Einsetzung von einer Funktion als Funktion von n Variablen an die Stelle der $(m+1)$ ten Variablen einer zweiten wird durch $(\Gamma_m \cdot B_mB_n)$ dargestellt.*

SATZ 3. *Sind die Substitutionsprozesse wie in den Sätzen 1 und 2 dargetellt, dann gestalten sie Ausdrücke der Form $(Yu_1u_2 \cdots u_n)$, wo Y eine eigentliche Kombination von Ordnung nicht zu gross ist, in andere Ausdrücke derselben Form um.*

Beweis: Für eine Umwandlung gilt

$$\begin{aligned} \vdash \Omega(Yu_1u_2 \cdots u_n) &= B_n\Omega Yu_1u_2 \cdots u_n \\ &= (B_{n-1}\Omega \cdot Y)u_1u_2 \cdots u_n. \end{aligned}$$

Für Zusammensetzungen: es sei $Z \equiv \Gamma_p \cdot B_pB_q$; dann

$$\begin{aligned} &\vdash Z(Xu_1u_2 \cdots u_m)(Yv_1v_2 \cdots v_n) \\ &\quad = (B_mZ \cdot BX)Iu_1u_2 \cdots u_m(Yv_1v_2 \cdots v_n) \\ &\quad = (\Gamma_m \cdot B_mB_n \cdot B_mZ \cdot BX)IYu_1u_2 \cdot u_mv_1v_2 \cdots v_n \\ &\quad = Uu_1u_2 \cdots u_mv_1v_2 \cdots v_m, \end{aligned}$$

wo U eigentlich ist, wenn nur X und Y eigentlich sind, und q gross genug ist, sodass $Yx_1x_2 \cdots x_{n+q}$ sich auf eine Kombination lauter Variablen reduziert. In der Tat reduziert $(Uu_1u_2 \cdots u_mv_1v_2 \cdots v_nx_1x_2 \cdots x_{p+q})$ sich

auf $(Xu_1u_2\cdots u_mx_1x_2\cdots x_p(Yv_1v_2\cdots v_nx_{p+1}\cdot x_{p+q}))$. Die Bedingung auf Y ist erfüllt, wenn wir es mit einem Substitutionsprozess zu tun haben.

SATZ 4. *X sei eine Kombination von Variablen und gewissen Etwasen $u_1, u_2, \cdots, u_m$. Dann gibt es einen Kombinator Y, sodass*

$$\vdash Yu_1u_2u_mx_1x_2\cdots x_n = X, \quad (u_i \textit{ als Variable behandelt}).$$

Gibt es weiter einen Kombinator Z, sodass

$$\vdash Z_1v_2v\cdots v_px_1x_2\cdots x_n = X \quad (v_i \textit{ als Variable betrachtet}),$$

wo $\vdash Iv_1v_2\cdots v_p = VIu_1u_2\cdots u_m, \quad (V \textit{ ein Kombinator}),$

(oder umgekehrt),

dann $\vdash Yu_1u_2\cdots u_m = Zv_1v_2\cdots v_p.$

*Beweis:** Wenn die $v_1v_2\cdots v_p$ dieselbe Reihe von Etwasen bildet, wie $u_1u_2, \cdots, u_m$, so folgt der Satz aus § 3, Satz 5. Sonst

$$\vdash Zv_1v_2\cdots v_p = (V\cdot BZ)Iu_1u_2\cdots u_p,$$

und $\vdash (V\cdot BZ)I = Y$ (§ 3, Satz 5).

$\therefore \quad \vdash Zv_1v_2\cdots v_p = Yu_1u_2\cdots u_m$ w. z. b. w.

* Der Beweis des ersten Teils des Satzes ist klar (§ 1, Satz 8).

A Test for the Type of Irrationality Represented by a Periodic Ternary Continued Fraction.

By J. B. Coleman.

1. *Introduction.* Let $(\overline{p_1, q_1; p_2, q_2; \cdots; p_k, q_k}; \cdots)$ denote a purely periodic ternary continued fraction,* of period $k \geqq 4$,† the partial quotient pairs being real numbers. Let D_1 denote the determinant

$$\begin{vmatrix} -p_1 & q_1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & (-1)^k & 0 \\ 1 & -p_2 & q_2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -p_3 & q_3 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -p_{k-1} & q_{k-1} & 1 \\ 0 & (-1)^k & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & -p_k & q_k \\ (-1)^k & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 \end{vmatrix}$$

and D_2 be the determinant derived from D_1 by replacing $(-1)^k$ by $(-1)^{k+1}$ where it occurs in the first row and in the second column, and replacing -1 by 1 in the last column.

In this paper I prove that the vanishing of D_1 or of D_2 is a necessary and sufficient condition for the reducibility of the characteristic equation, when p_i and q_i are rational integers, in which case the ternary continued fraction represents a rational number or a quadratic irrationality.‡ If p_i and q_i are any real numbers, the vanishing of D_1 or of D_2 is a sufficient condition for the reducibility of the characteristic equation.

We proceed to prove the above statements by first finding determinant forms for the convergents and other expressions involved in the characteristic equation. This is done in sections 2-6. In sections 7-11, following the general proof, are some corollaries and numerical examples.

2. Let three sequences satisfying the recursion formula

$$W_n = q_n W_{n-1} + p_n W_{n-2} + W_{n-3} \tag{1}$$

* References for previous history: C. G. J. Jacobi, *Werke*, Vol. 6, p. 385; O. Perron, *Mathematische Annalen*, Vol. 64, p. 1; D. N. Lehmer, *Proceedings of the National Academy of Sciences*, Vol. 4, p. 360; H. P. Daus, *American Journal of Mathematics*, Vol. 51, p. 67; O. Perron, "Die Lehre von den Kettenbrüchen."

† If $k < 4$ the expressions for D_1 and D_2 must be interpreted as shown in section 9.

‡ For convergence conditions see O. Perron, article cited.

with the initial values

$$\text{(1a)} \qquad (0,0,1) \quad (0,1,0) \quad (1,0,0)$$

be denoted, respectively, by C_n, B_n, A_n.

The characteristic equation is

$$\text{(2)} \qquad \rho^3 - M\rho^2 + N\rho - 1 = 0,$$

in which

$$\text{(3a)} \qquad M = A_{k-2} + B_{k-1} + C_k,$$

$$\text{(3b)} \qquad N = (A_{k-2}, B_{k-1}) + (A_{k-2}, C_k) + (B_{k-1}, C_k).$$

$(B_{k-1}, C_k) = B_{k-1}C_k - B_kC_{k-1}$, &c.

Since by a theorem of Lehmer's* the roots, $\sigma_{1,1}$, and $\sigma_{2,1}$, of the cubic equation representing the expansion are related to ρ_1, the principal root of (2), as follows;

$$\text{(4a)} \qquad \sigma_{1,1} = (B_k\rho_1 + A_kB_{k-2} - A_{k-2}B_k) / (A_k\rho_1 + A_{k-1}B_k - A_kB_{k-1})$$

$$\text{(4b)} \qquad \sigma_{2,1} = (C_{k-1}\rho_1 + C_{k-2}A_{k-1} - C_{k-1}A_{k-2}) / (A_{k-1}\rho_1 + A_kC_{k-1} - A_{k-1}C_k),$$

the rationality or type of irrationality represented by the continued fraction may be determined from a discussion of the characteristic equation. Since (2) is of the third degree, if it is reducible when M and N are rational integers, it must have a factor $\rho - 1$ or $\rho + 1$. Hence the necessary and sufficient conditions for reducibility under these conditions, are

$$\text{(5a)} \qquad -M + N = 0 \quad \text{or}$$

$$\text{(5b)} \qquad M + N + 2 = 0.$$

If M and N are any real numbers, conditions (5a) or (5b) will be sufficient to insure reducibility.

3. To find determinant forms for A_n, B_n and C_n.

Let
$$M\begin{pmatrix} a_1, a_2, a_3, \cdots, a_n \\ \quad b_2, b_3, \cdots, b_n \end{pmatrix}_n$$

denote the determinant

$$\text{(6)} \qquad \begin{vmatrix} a_1 & 1 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ -b_2 & a_2 & 1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ 1 & -b_3 & a_3 & 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 & -b_{n-1} & a_{n-1} & 1 \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & 1 & -b_n & a_n \end{vmatrix}.$$

* D. N. Lehmer, *loc. cit.*

Following the usage in ordinary continued fractions we shall call (6) a *continuant.*

Expanding (6) according to the elements of its last row,

$$(7a)\qquad M\begin{pmatrix} a_1, a_2, a_3, \cdots, a_n \\ b_2, b_3, b_4, \cdots, b_n \end{pmatrix}_n = a_n M\begin{pmatrix} a_1, a_2, \cdots, a_{n-1} \\ b_2, b_3, \cdots, b_{n-1} \end{pmatrix}_{n-1} + b_n M\begin{pmatrix} a_1, a_2, \cdots, a_{n-2} \\ b_2, \cdots, b_{n-2} \end{pmatrix}_{n-2} + M\begin{pmatrix} a_1, a_2, \cdots, a_{n-3} \\ b_2, b_3, \cdots, b_{n-3} \end{pmatrix}_{n-3}, \qquad (n > 3).$$

By (1) and (1a) we have directly

$$(7b)\qquad A_2 = M(q_2), \quad A_3 = M\begin{pmatrix} q_2, q_3 \\ p_3 \end{pmatrix}_2, \quad A_4 = M\begin{pmatrix} q_2, q_3, q_4 \\ p_3, p_4 \end{pmatrix}_3.$$

From (7a) and (1) we have immediately

$$(8)\qquad A_n = M\begin{pmatrix} q_2, q_3, q_4, \cdots, q_n \\ p_3, p_4, \cdots, p_n \end{pmatrix}_{n-1}, \qquad (n = 2, 3, \cdots, n),$$

for any n, provided it is true for three successively lower values of n. But by (7b) the relation (8) is true for $n = 2, 3, 4$; hence, by induction, it is true for all values of n.

In the same way we find *

$$(9)\qquad B_n = p_1 A_n + M\begin{pmatrix} q_3, q_4, \cdots, q_n \\ p_4, p_5, \cdots, p_n \end{pmatrix}_{n-2}, \qquad (n = 3, 4, \cdots).$$

$$(10)\qquad C_n = M\begin{pmatrix} q_1, q_2, q_3, \cdots, q_n \\ p_2, p_3, \cdots, p_n \end{pmatrix}_n, \qquad (n = 1, 2, 3, \cdots).$$

4. By the recursion formulae (1) and (1a) it is found that

$$(A_{-2}, B_{-1}) = 1, \quad (A_{-1}, B_0) = 0, \quad (A_0, B_1) = 0, \quad (A_1, B_2) = 1,$$

and

$$(A_n, B_{n+1}) = -p_{n+1}(A_{n-1}, B_n) - q_n(A_{n-2}, B_{n-1}) + (A_{n-3}, B_{n-2}), \qquad (n = 1, 2, 3, \cdots).$$

By direct calculation it is found that

$$(A_2, B_3) = M(-p_3)_1, \quad (A_3, B_4) = M\begin{pmatrix} -p_3, -p_4 \\ -q_3 \end{pmatrix}_2,$$

$$(A_4, B_5) = M\begin{pmatrix} -p_3, -p_4, -p_5 \\ -q_3, -q_4 \end{pmatrix}_3.$$

Hence

$$(11)\qquad (A_n, B_{n+1}) = M\begin{pmatrix} -p_3, -p_4, \cdots, -p_{n+1} \\ -q_3, -q_4, \cdots, -q_n \end{pmatrix}_{n-1}, \qquad (n = 2, 3, \cdots),$$

by inductive reasoning similar to that employed in deriving (8).

* A set of continuants similar to (8), (9) and (10) may be written for the convergents involved in quaternary, or in n-ary, continued fractions. In every case the proof, by induction, involves assuming that the convergent is represented by its continuant for n successive orders and proving it true for the next higher order.

By (1) and by definition

$$(A_n, C_{n+2}) = q_{n+2}(A_n, C_{n+1}) - (A_{n-1}, C_n).$$

Hence

$$(A_n, C_{n+2}) = q_{n+2} M \begin{pmatrix} -p_2, -p_3, \cdots, -p_{n+1} \\ -q_2, \cdots, -q_n \end{pmatrix}_n - M \begin{pmatrix} -p_2, -p_3, \cdots, -p_n \\ -q_2, \cdots, -q_{n-1} \end{pmatrix}_n$$

since by a process similar to that used in deriving (11) we have

$$(12) \qquad (A_n, C_{n+1}) = M \begin{pmatrix} -p_2, -p_3, -p_4, \cdots, -p_{n+1} \\ -q_2, -q_3, \cdots, -q_n \end{pmatrix}_n, \quad (n = 1, 2, \cdots).$$

Also from similar considerations

$$(13) \qquad (B_n, C_{n+1}) = M \begin{pmatrix} -p_1, -p_2, \cdots, -p_{n+1} \\ -q_1, -q_2, \cdots, -q_n \end{pmatrix}_{n+1}, \qquad (n = 0, 1, 2, \cdots).$$

5. Proof that $D_2 = M + N + 2$.

Expanding D_2 in terms of the elements in the last row we obtain two determinants of order k. We shall designate by D_3 the minor of 1, and by D_4 the minor of $(-1)^k$, in the last row. Next we expand D_3 in terms of the elements of the last row and last column, by Cauchy's method. Applying (13) three times, (8) twice, and (9) three times to terms of this expansion, we obtain

$$(14) \qquad D_3 = -p_k(B_{k-2}, C_{k-1}) - q_{k-1}(B_{k-3}, C_{k-2}) + (B_{k-4}, C_{k-3}) \\ + q_{k-1}B_{k-2} + p_{k-1}B_{k-3} + B_{k-4} + (A_{k-2}, B_{k-1}) + 1.$$

Now by the recursion formulae (1) for B_n and C_n and by definition, (B_{k-1}, C_k) reduces to $-p_k(B_{k-2}, C_{k-1}) - (B_{k-3}, C_{k-1})$ and also (B_{k-3}, C_{k-1}) reduces to $q_{k-1}(B_{k-3}, C_{k-2}) - (B_{k-4}, C_{k-3})$. Hence the first three terms of (14) reduce to (B_{k-1}, C_k). By (I) for B_n the next three terms of (14) reduce to B_{k-1}. Substituting these terms in (14) it becomes

$$(15) \qquad D_3 = (B_{k-1}, C_k) + B_{k-1} + (A_{k-2}, B_{k-1}) + 1.$$

Next expand D_4 in terms of the elements of the last row and next to the last column, by Cauchy's method. Applying (12) twice, (10) five times and (8) once to the expansion gives

$$(16) \qquad D_4 = q_k(A_{k-2}, C_{k-1}) - (A_{k-3}, C_{k-2}) + p_k C_{k-2} + q_{k-1}q_k C_{k-2} \\ + q_k p_{k-1} C_{k-3} + q_k C_{k-4} + C_{k-3} + A_{k-2} + 1.$$

By recursion formulae (1) for A_n and C_n the first two terms of (16) reduce to (A_{k-2}, C_k). By (1) for C_k and C_{k-1} the next five terms become C_k. Substituting these values in (16) it becomes

(17) $$D_4 = (A_{k-2}, C_k) + C_k + A_{k-2} + 1.$$

Combining (15) and (17) gives

$$D_2 = D_3 + D_4 = C_k + B_{k-1} + A_{k-2} + (B_{k-1}, C_k) + (A_{k-2}, C_k) + (A_{k-2}, B_{k-1}) + 2.$$

Hence by (3) and (4),

$$D_2 = M + N + 2.$$

6. To show that

(18) $$D_1 = C_k + B_{k-1} + A_{k-2} - (B_{k-1}, C_k) - (A_{k-2}, C_k) - (A_{k-2}, B_{k-1})$$

it is not necessary to expand it completely as was done for D_2. The elements of the two determinants correspond except that three elements of each are replaced in the other by the same elements with their signs changed. Hence the expansion of D_1 may be obtained from that of D_2 by making appropriate changes of sign. The result of making these changes of sign in the preceding section produces (18). Thus by (3) and (4), $D_1 = M - N$.

From (5) and (6) it is now evident that the proof of the original statements is complete, i. e., that for the reducibility of the characteristic equation of a periodic ternary (continued) fraction, the vanishing of D_1 or D_2 is, (a) a necessary and sufficient condition where p_i and q_i are rational integers, (b) a sufficient condition when p_i and q_i are any real numbers.

7. Since D_1 and D_2 are linear in any particular p_i and q_i, it is obvious that if $k-1$ pairs of partial quotients be selected arbitrarily, the remaining pair may be selected in an infinite number of ways so as to make the characteristic equation reducible, either by the root 1 or -1.

In the same way D_1 and D_2 are linear in any q_i and p_{i+1}, so that the same statement may be made for such a pair as for the pair p_i and q_i.

8. Below are listed six general conditions under which the characteristic equation will be reducible. The classification is made according to the method of derivation from D_1 or D_2.

A. $p_i = -q_i$.

B. $p_i = q_i + 2$.

C. $p_1 = q_k = 0,\ p_i = -q_{i-1}, \qquad (i = 2, 3, \cdots, k)$.

D. $p_1 = 0,\ q_k = 2,\ p_2 + q_1 + 2 = 0,\ p_k + q_{k-1} = 2,\ p_i = -q_{i-1}, \qquad (i = 3, \cdots, k-1)$.

E_1. k being odd, $p_1 = q_k = 0,\ p_2 = q_1,\ p_k = q_{k-1},\ p_i = q_{i-1} + 2, \qquad (i = 3, \cdots, k-1)$.

E_2. k being even, $p_1 = 2,\ q_k = 0,\ p_i = q_{i-1} + 2, \qquad (i = 2, 3, \cdots, k)$.

F_1. k odd, $p_1 = 0$, $q_k = -2$, $p_i = q_{i-1} + 2$, $\qquad (i = 2, 3, \cdots, k)$.
F_2. k even, $p_1 = 2$, $q_k = -2$, $p_2 = q_1$, $p_k = q_{k-1}$, $p_i = q_{i-1} + 2$,
$(i = 3, \cdots, k-1)$.

Two special cases arise amongst these. For $k = 2$, condition D becomes $p_1 = 0$, $q_2 = 2$, $q_1 = -p_2$. Also for $k = 2$, F_2 becomes $p_1 = 2$, $q_2 = -2$, $p_2 = q_1 - 2$.

Reducibility under A results from the fact that under this condition the sum of the odd columns in D_1 is equal to the sum of the even numbered columns.

Reducibility under condition B may be shown from the fact that the sum of the elements in the i-th row of D_1 is $-p_i + q_i + 2$ when k is even, and the same is true for the i-th row of D_2 when k is odd. Hence under this condition the root of the characteristic equation will be $(-1)^k$. This condition was found and proved by Lehmer.

Conditions C, D, E, F result from a consideration of the sums and differences of the two sets of alternate rows of D_1 and D_2.

9. The vanishing of D_1 or D_2, in the special cases where $k = 1, 2, 3$, may be obtained from the general formulae by observing the following. It is necessary to have the elements involving powers of -1 always occupy the three positions indicated below;

Row	Column
1	k
k	2
$k+1$	1

In case $k = 1, 2, 3$ these elements are to be added to any other elements which may occupy the same position in the determinant.

10. The character of the roots of a reducible characteristic equation when M and N are rational integers.

A. For $D_1 = 0$, or $M - N = 0$, one root of the characteristic equation is 1. The other two roots will also be rational only in case $M = N = 3$ or $M = N = -1$. When $-1 < M = N < 3$ the other roots will be imaginary. When $M = N < -1$ or $M = N > 3$, the two remaining roots will be quadratic irrationalities.

B. For $D_2 = M + N + 2 = 0$, one root of the characteristic equation is -1 and the other two are always quadratic irrationalities.

11. Numerical examples in which the characteristic equation is reducible.

A. An example in which the partial quotients are positive integers and for which D_1 vanishes.

Given $(\overline{2, 2;\ 4, 1;\ 8, 1;\ 8, 5;}\ \cdots)$, k, the number of pairs per period being 4.

The characteristic equation is $\rho^3 - 185\rho^2 + 185\rho - 1 = 0$. The principal root, $\rho_1 = 92 + \sqrt{8463}$.

By (4a),
$$\sigma_{1,1} = \frac{11163 + 121\sqrt{8463}}{4905 + 54\sqrt{8463}}$$

and by (4b),
$$\sigma_{2,1} = \frac{2147 + 23\sqrt{8463}}{285 + 9\sqrt{8463}}.$$

B. An example containing positive and negative partial quotients, D_1 vanishing for the set.

Given $(\overline{2, 3;\ 3, -1;\ 2, 4;}\ \cdots)$, k being 3.

The characteristic equation is $\rho^3 - 7\rho^2 + 7\rho - 1 = 0$.

The principal root is $\rho_1 = 3 + 2\sqrt{2}$.

By (4a), $\sigma_{1,1} = (2 - \sqrt{2})/2$,

and by (4b), $\sigma_{2,1} = -(6 + 3\sqrt{2})/4$.

C. An example of the same type as *B*.

Given $(\overline{2, 3;\ 1, -1;}\ \cdots)$, k being 2.

The characteristic equation is $\rho^3 - 1 = 0$. The general conditions for convergence are not satisfied, so that the expansion does not give a limit for B_n/A_n or C_n/A_n.

D. An example involving positive and negative partial quotients, D_1 and D_2 both vanishing for the set.

Given $(\overline{-1, 1;\ 3, -3;}\ \cdots)$, k being two.

The characteristic equation is $\rho^3 - \rho^2 + \rho - 1 = 0$.

The principal root $\rho_1 = -1$.

By (4a), $\sigma_{1,1} = -1$,

and by (4b), $\sigma_{2,1} = 0$.

E. An example involving fractional partial quotients, D_1 vanishing for the set.

Given $(\overline{1/2, 2/3;\ 1/3, -4/5;}\ \cdots)$, k being 2.

The characteristic equation is $\rho^3 - 3/10\rho^2 + 3/10\rho - 1 = 0$.

This equation is reducible but again the conditions for convergence are not satisfied.

F. An example involving irrational partial quotients, D_1 vanishing for the set.*

Given $(\overline{\sqrt{2}, \sqrt{3}; \sqrt{6}, \sqrt{2}; -\sqrt{3}, \sqrt{6}}; \cdots)$, k being 3.

The characteristic equation is $\rho^3 - 14\rho^2 + 14\rho - 1 = 0$.

The principal root is $\rho_1 = (13 + \sqrt{165})/2$.

By (4a), $$\sigma_{1,1} = \sqrt{2}(\ 5 + \sqrt{165})/10,$$

and by (4b), $$\sigma_{2,1} = \sqrt{3}(15 + \sqrt{165})/10.$$

* Under the given conditions the characteristic equation is reducible, even when irrational partial quotients are involved. However, for such a set of partial quotients, the type of irrationality defined by the continued fraction will not in general be quadratic, as was the case in example *F*.

On the Separation Property of the Roots of the Secular Equation.

By E. T. Browne.

1. *Introduction.* Let A be any square matrix, real or complex, of order n. If I is the unit matrix, $A - \lambda I$ is called the *characteristic matrix* of A; the determinant of the characteristic matrix is called the *characteristic determinant* of A; the equation obtained by equating this determinant to zero is called the *characteristic equation* of A; and the roots of this equation are called the *characteristic roots* of A. In particular, if A is real and symmetric, i. e., $a_{ij} = a_{ji}$, the characteristic equation is of great importance and is called the *secular equation* since it was first used by Laplace in the determination of the secular inequalities of the planets.

The secular equation has been widely studied and many beautiful properties of it have been discovered. For example, let us following Weber * denote by $L_i(\lambda)$ the determinant of order i standing in the upper left hand corner of $A - \lambda I$. Weber gives a proof that the roots of $L_i(\lambda) = 0$ are all real and are *separated* by the roots of $L_{i-1}(\lambda) = 0$. However, it may happen that a root ρ of multiplicity m of the latter equation is a root of multiplicity $m - 1$, m or even $m + 1$ of the former, so that if $L_{i-1}(\lambda) = 0$ has a multiple root the sense in which the previously mentioned "separation" takes place is not exactly clear. It is the purpose of this paper to study this separation property. In doing so we shall employ merely the simplest properties of algebraic equations together with a well known theorem which in the study of the characteristic equation of a matrix is one of the most useful with which the author is acquainted; viz.,

If A is a Hermitian (real symmetric) matrix of order n, there exists a unitary (real orthogonal) matrix R such that $\overline{R}'AR = N$ $(R'AR = N)$, where N has as elements in its main diagonal the (real) characteristic roots of A and zeros elsewhere.†

This theorem was used by Bromwich ‡ in his proof that if $\alpha + i\beta$ is a characteristic root of a matrix A whose elements are real or complex, and if $\rho_1 \leqq \cdots \leqq \rho_n$ are the characteristic roots of $(A + \bar{A}')/2$ and $i\mu_1, \cdots, i\mu_n$ are

* Weber, *Lehrbuch der Algebra*, Braunschweig (1898), Vol. I, pp. 307-311.

† Dickson, *Modern Algebraic Theories*, Chicago (1926), pp. 74-76; Kowalewski, *Determinantentheorie*, Berlin (1925), pp. 194-198.

‡ Bromwich, "On the Roots of the Characteristic Equation of a Linear Substitution," *Acta Mathematica*, Vol. 30 (1906), pp. 295-304.

the characteristic roots of $(A-\bar{A}')/2$, then $\rho_1 \leqq \alpha \leqq \rho_n$ and $|\beta|$ does not exceed the greatest of $|\mu_1|, \cdots, |\mu_n|$. The same theorem was employed by the author* in the proof that if λ is a charateristic root of a matrix A and if G and s are respectively the largest and smallest characteristic roots of $A\bar{A}'$, then $s \leqq \lambda\bar{\lambda} \leqq G$.

2. *Transformation of a Hermitian Matrix.* Let us suppose then that A is a Hermitian matrix of order n. Denote by A_τ the principal minor matrix of order τ standing in the upper left hand corner of A and by $L_\tau(\lambda)=0$ the characteristic equation of A_τ. If $\rho_1 \leqq \cdots \leqq \rho_\tau$ are the characteristic roots of A_τ there exists a unitary matrix $P=(p_{ij})$ such that $\bar{P}'A_\tau P = B_\tau$, where B_τ has as elements in its main diagonal the roots $\rho_1, \cdots, \rho_\tau$ and zeros elsewhere. If $A_{\tau+1}$ be the Hermitian matrix of order $\tau+1$ formed by adjoining to A_τ an additional row $x_1, \cdots, x_\tau, x_{\tau+1}$ ($x_{\tau+1}$ real), and a column consisting of the conjugates of these elements, and if R be the unitary matrix

$$R = \begin{pmatrix} P, & 0 \\ 0, & 1 \end{pmatrix}$$

formed by adjoining to P an additional row and column consisting entirely of zeros except in the last place, it is easy to verify that $\bar{R}'A_{\tau+1}R = B_{\tau+1}$, $B_{\tau+1}$ being a matrix of the form:

$$B_{\tau+1} = \begin{pmatrix} \rho_1, \cdots, 0, \bar{X}_\tau \\ \cdots\cdots\cdots \\ 0, \cdots, \rho_\tau, \bar{X}_1 \\ X_1, \cdots, X_\tau, X_{\tau+1} \end{pmatrix} \tag{1}$$

where

$$X_{\tau+1} = x_{\tau+1}; \quad X_j = \sum_{i=1}^{\tau} p_{ij}x_i \qquad (j=1, \cdots, \tau). \tag{2}$$

Under such a transformation the characteristic equations $L_\tau(\lambda)=0$ and $L_{\tau+1}(\lambda)=0$ of A_τ and $A_{\tau+1}$ are unaltered.

Expanding the characteristic determinant of $B_{\tau+1}$ according to the elements of its last row and last column, $L_{\tau+1}(\lambda)$ may be written

$$L_{\tau+1}(\lambda) = -\sum_{i=1}^{\tau} X_i\bar{X}_iR_i(\lambda) + (x_{\tau+1}-\lambda)L_\tau(\lambda), \tag{3}$$

where the $R_i(\lambda)$ are defined by the relations

$$(\rho_i-\lambda)R_i(\lambda) = (\rho_1-\lambda)\cdots(\rho_\tau-\lambda) = L_\tau(\lambda) \tag{4}$$

and are therefore real. Manifestly $R_i(\rho_j)=0$ $(i \neq j)$ while if the ρ's are all distinct $R_i(\rho_i) \neq 0$.

* "The Characteristic Equation of a Matrix," *Bulletin of the American Mathematical Society*, Vol. 34 (1928), pp. 363-368.

3. *The Vanishing of Certain X's.* Since P is nonsingular $(X_1,\cdots,X_\tau)=(0,\cdots,0)$ if, and only if, $(x_1,\cdots,x_\tau)=(0,\cdots,0)$. Let us suppose then that the X's are not all zero but that $X_{\gamma+1},\cdots,X_{\gamma+m}$ which correspond in (1) to a root $\rho_{\gamma+1}=\cdots=\rho_{\gamma+m}=\rho$ (say) of multiplicity m of A_τ are all zero. We then have

$$\sum_{i=1}^{\tau} p_{ij}x_i=0 \qquad (j=\gamma+1,\cdots,\gamma+m) \tag{5}$$

so that the set $(x_1,\cdots,x_\tau)$ is a solution of the system of homogeneous linear equations (5) whose coefficients are the $(\gamma+1)$th, $\cdots$, $(\gamma+m)$th columns of P. But from the manner in which P was built up * we have also the following

$$\sum_{j=1}^{\tau}(a_{ij}-\rho\delta_{ij})p_{jr}=0 \qquad (i=1,\cdots,\tau),\ (r=\gamma+1,\cdots,\gamma+m) \tag{6}$$

where δ_{ij} is the Kronecker symbol and is equal to 1 if $i=j$; 0 if $i\neq j$. Thus the $\tau-m$ linearly independent rows of $A_\tau-\rho I$ are also solutions of (5), and since the latter system has at most $\tau-m$ linearly independent solutions, it follows that the set $(x_1,\cdots,x_\tau)$ depends linearly on the rows of $A_\tau-\rho I$. Conversely, if $(x_1,\cdots,x_\tau)$ depends linearly on the rows of $A_\tau-\rho I$, $X_{\gamma+1}=\cdots=X_{\gamma+m}=0$. We therefore have the following theorem:

THEOREM I. *If ρ is a characteristic root of multiplicity m of A_τ and if $X_{\gamma+1},\cdots,X_{\gamma+m}$ are X's corresponding to this multiple root in the matrix $B_{\tau+1}$; then $X_{\gamma+1}=\cdots=X_{\gamma+m}=0$, if, and only if, the bordering set $x_1,\cdots,x_\tau$ depends linearly on the rows of $A_\tau-\rho I$.*

4. *The Separation Property.* Let us now suppose that in (3) all the ρ's are distinct and none of the X's is zero. Since $L_{\tau+1}(\lambda)=(-\lambda)^{\tau+1}+\cdots$, manifestly $L_{\tau+1}(-\infty)>0$ whether τ is even or odd. Also

$$\begin{aligned} L_{\tau+1}(\rho_i)&=-X_i\bar{X}_iR_i(\rho_i)\\ &=-X_i\bar{X}_i(\rho_1-\rho_i)\cdots(\rho_{i-1}-\rho_i)(\rho_{i+1}-\rho_i)\cdots(\rho_\tau-\rho_i)\\ &=(-1)^i k_i \quad \text{where } k_i>0 \qquad (i=1,\cdots,\tau). \end{aligned} \tag{7}$$

Further, $L_{\tau+1}(\infty)$ has the same sign as $(-1)^{\tau+1}$. Using for uniformity the notation $\rho_0=-\infty$, $\rho_{\tau+1}=\infty$, we may say that (7) holds also for $i=0$ and $i=\tau+1$. It is clear then that in each of the open intervals (ρ_{i-1},ρ_i) $(i=1,\cdots,\tau+1)$ there is exactly one root σ_i of $L_{\tau+1}(\lambda)=0$. We therefore have

* Kowalewski, *loc. cit.*, pp. 195-196.

THEOREM II. *If the characteristic roots $\rho_1, \cdots, \rho_\tau$ of a Hermitian matrix A_τ are all distinct and if $A_{\tau+1}$ is the Hermitian matrix formed by adjoining to A_τ a row $x_1, \cdots, x_\tau, x_{\tau+1}$ ($x_{\tau+1}$ real)* and *a column $\bar{x}_1, \cdots, \bar{x}_\tau, x_{\tau+1}$, then if the set $x_1, \cdots, x_\tau$ does not depend linearly on the rows of any of the matrices $A_\tau - \rho_i I$ $(i = 1, \cdots, \tau)$ the $\tau + 1$ characteristic roots $\sigma_1, \cdots, \sigma_{\tau+1}$ of $A_{\tau+1}$ are all distinct and are separated by the ρ's.*

Suppose, however, that $\rho_{\gamma+1} = \cdots = \rho_{\gamma+e} = \rho$ is a root of multiplicity e of $L_\tau(\lambda) = 0$. Then evidently each $R_i(\lambda)$ is divisible by $(\rho - \lambda)^{e-1}$ while $R_i(\lambda)$ $(i = \gamma + 1, \cdots, \gamma + e)$ are not divisible by $(\rho - \lambda)^e$. Writing $R_i(\lambda) = (\rho - \lambda)^{e-1} S_i(\lambda)$ and noting that

$$S_{\gamma+1}(\lambda) = \cdots = S_{\gamma+e}(\lambda) = S_\rho(\lambda), \quad \text{say,}$$

we may write $L_{\tau+1}(\lambda)$ in the form

$$L_{\tau+1}(\lambda) = (\rho - \lambda)^{e-1} f(\lambda)$$

where $f(\lambda)$ is an expression of the type (3) with the root ρ now playing the role of a simple root and with the coefficient $X_\rho \bar{X}_\rho$ of $S_\rho(\lambda)$ satisfying the condition

$$X_\rho \bar{X}_\rho = \sum_{i=\gamma+1}^{\gamma+e} X_i \bar{X}_i.$$

Evidently $X_\rho = 0$ if, and only if, the set $x_1, \cdots, x_\tau$ depends linearly on the rows of $A_\tau - \rho I$.

If now $L_\tau(\lambda) = 0$ has the $m \leqq \tau$ distinct roots $\rho_1 < \cdots < \rho_m$ of multiplicities $e_1, \cdots, e_m$, respectively, we may proceed with regard to each of these roots as we did with regard to ρ until finally $L_{\tau+1}(\lambda)$ may be written in the form

$$L_{\tau+1}(\lambda) = (\rho_1 - \lambda)^{e_1-1} \cdots (\rho_m - \lambda)^{e_m-1} F(\lambda)$$

where $F(\lambda)$ is an expression of the type (3) with each ρ playing the role of a simple root. If the set $x_1, \cdots, x_\tau$ does not depend linearly on the rows of any of the matrices $A_\tau - \rho_i I$ all of the X's entering $F(\lambda)$ are different from zero. Hence, writing $\rho_0 = -\infty$, $\rho_{m+1} = \infty$ it follows that $F(\lambda) = 0$ has exactly one root in each of the open intervals (ρ_{i-1}, ρ_i) $(i = 1, \cdots, m + 1)$.

We have therefore proved

THEOREM III. *If $L_\tau(\lambda) = 0$ has the m distinct roots $\rho_1 < \cdots < \rho_m$ of multiplicities $e_1, \cdots, e_m$, respectively, and if the bordering set $x_1, \cdots, x_\tau$ does not depend linearly on the rows of any of the matrices $A_\tau - \rho_i I$ $(i = 1, \cdots, m)$, then each ρ_i is a root of $L_{\tau+1}(\lambda) = 0$ of multiplicity exactly $e_i - 1$, while in each of the open intervals (ρ_{i-1}, ρ_i) $(i = 1, \cdots, m + 1)$ there lies exactly one root of $L_{\tau+1}(\lambda) = 0$.*

Suppose now that $\rho_{\gamma+1} = \cdots = \rho_{\gamma+m} = \rho$ is a root of multiplicity m of $L_\tau(\lambda) = 0$ and that the set $x_1, \cdots, x_\tau$ depends linearly on the rows of $A_\tau - \rho I$, so that $(X_{\gamma+1}, \cdots, X_{\gamma+m}) = (0, \cdots, 0)$. From the determinantal form of $L_{\tau+1}(\lambda)$ it is manifest that the latter contains $(\rho - \lambda)^m$ as a factor, so that ρ is a root of $L_{\tau+1}(\lambda) = 0$ of multiplicity at least m. Indeed, if

$$x_j = \sum_{i=1}^{\tau} (a_{ij} - \rho\delta_{ij}) c_i = \sum_{i=1}^{\tau} a_{ij} c_i - \rho c_j$$

it follows from an examination of the rank of $B_{\tau+1}$ that ρ will be a root of multiplicity $m + 1$ or m of $L_{\tau+1}(\lambda) = 0$ according as $x_{\tau+1}$ is or is not equal to

$$\rho(1 - \sum_{j=1}^{\tau} c_j \bar{c}_j) + \sum_{i,j}^{1,\ldots,\tau} a_{ij} c_i \bar{c}_j.$$

Hence we have the following theorem:

THEOREM IV. *A root ρ of multiplicity m of $L_\tau(\lambda) = 0$ will be a root of multiplicity at least m (and at most $m + 1$) of $L_{\tau+1}(\lambda) = 0$ if, and only if, the bordering set $x_1, \cdots, x_\tau$ depends linearly on the rows of $A_\tau - \rho I$.*

5. *Number of Negative and of Positive Roots of $A_{\tau+1}$.* Let the ν distinct negative roots of $L_\tau(\lambda) = 0$ be $\rho_1 < \cdots < \rho_\nu$ of multiplicities $e_1, e_2, \cdots, e_\nu$, respectively. If the X's corresponding to the roots $\rho_{i_1}, \cdots, \rho_{i\gamma}$ are all zero the latter are roots of $L_{\tau+1}(\lambda) = 0$ of multiplicities at least $e_{i_1}, \cdots, e_{i\gamma}$, respectively. If for the remaining ρ's, $\rho_{i\gamma_{+1}}, \cdots, \rho_{i_\nu}$ the corresponding X's are not all zero, these are roots of $L_{\tau+1}(\lambda) = 0$ of multiplicities exactly $e_{i\gamma_{+1}} - 1, \cdots, e_{i_\nu} - 1$, respectively, while in each of the $\nu - \gamma$ open intervals

(8) $$-\infty, \quad \rho_{i\gamma_{+1}} \cdots, \rho_{i_\nu},$$

there is exactly one (negative) root of $L_{\tau+1}(\lambda) = 0$. Hence, the latter equation has at least as many negative roots as $L_\tau(\lambda) = 0$. Similarly, $L_{\tau+1}(\lambda) = 0$ has at least as many positive roots as $L_\tau(\lambda) = 0$.

If zero is a root of multiplicity m of $L_\tau(\lambda) = 0$ and is likewise a root of multiplicity at least m of $L_{\tau+1}(\lambda) = 0$, it is clear that the latter equation can have at most one more negative (positive) root than the former; while if zero is a root of multiplicity exactly $m - 1$ of $L_{\tau+1}(\lambda) = 0$ by adjoining 0 to the sequence (8) it follows that the latter equation has exactly one more negative root and likewise one more positive root than $L_\tau(\lambda) = 0$.

We may state the theorem as follows:

THEOREM V. *If m, ν and μ represent the numbers of zero, negative and positive roots of $L_\tau(\lambda) = 0$ and if Z, N and P represent the corresponding numbers for $L_{\tau+1}(\lambda) = 0$, then, if $Z = m - 1$, $N = \nu + 1$* and *$P = \mu + 1$;*

if $Z=m$, $N=\nu+1$ or ν, $P=\mu$ *or* $\mu+1$; *and finally, if* $Z=m+1$, $N=\nu$ *and* $P=\mu$.

6. *The Signature of a Hermitian Matrix.* If $L_\tau(\lambda)=0$ has ν negative roots and μ positive roots, the difference $\mu-\nu$ is called the *signature** of A_τ. Denote by M_i the determinant of the matrix A_i. Suppose now that A_τ is non-singular, i. e., $M_\tau\neq 0$. If $A_{\tau+1}$ is also non-singular, by Theorem V $L_{\tau+1}(\lambda)=0$ will have ν negative and $\mu+1$ positive roots or $\nu+1$ negative and μ positive roots according as M_τ and $M_{\tau+1}$ have the same sign or opposite signs. That is, the signature of $A_{\tau+1}$ is greater or less by one than the signature of A_τ according as the sequence of two terms M_τ, $M_{\tau+1}$ presents a permanence or a variation of sign.

But if $A_{\tau+1}$ is singular and therefore $L_{\tau+1}(\lambda)=0$ has one zero root, the latter has exactly ν negative and μ positive roots. If further $A_{\tau+2}$ is non-singular, $L_{\tau+2}(\lambda)=0$ has by Theorem V exactly $\nu+1$ negative and $\mu+1$ positive roots. Hence, $M_{\tau+2}$ is of opposite sign to M_τ. Moreover, the signatures of $A_{\tau+2}$ and A_τ are the same. Noting that the matrix $\begin{pmatrix} 0, & a_{12} \\ \bar{a}_{12}, & 0 \end{pmatrix}$ $(a_{12}\neq 0)$ has one negative and one positive characteristic root, it is clear that we have established Gundelfinger's † rule for determining the signature of a regularly arranged Hermitian or real symmetric matrix.

THEOREM VI. *If a Hermitian or a real symmetric matrix of rank r is regularly arranged, i. e., if the rows and columns are so arranged that no two consecutive terms in the sequence*

$$M_0=1,\quad M_1=a_{11},\cdots,M_r=\begin{vmatrix} a_{11},\cdots,a_{1r} \\ \cdots\cdots \\ a_{r1},\cdots,a_{rr}\end{vmatrix},\tag{9}$$

are zero and $M_r\neq 0$, *the signature of the matrix is equal to the difference between the number of permanences of sign and the number of variations of sign in the sequence* (9), *where a vanishing term may be counted as either positive or negative, but must be counted.*

7. *Application to Hermitian Matrices which are not Regularly Arranged.* Suppose now that both $A_{\tau+1}$ and $A_{\tau+2}$ are singular while A_τ is not. Let us denote by Z, N and P the numbers of zero, negative and positive roots of an equation under consideration. It is clear that if for $L_\tau(\lambda)=0$ we have

$$L_\tau(\lambda):\qquad Z=0,\quad N=\nu,\quad P=\mu,$$

* μ is sometimes called the index of A; cf. Dickson, *loc. cit.*, p. 71.

† Gundelfinger, "Zur Theorie der quadratischen Formen," *Crelle*, Vol. 91 (1881), p. 225; cf. also Dickson, *loc. cit.*, pp. 87-88.

then for $L_{\tau+1}(\lambda)=0$ we have

$$L_{\tau+1}(\lambda): \quad Z=1, \quad N=\nu, \quad P=\mu,$$

and for $L_{\tau+2}(\lambda)=0$ we have one of the following

$$(10) \qquad L_{\tau+2}(\lambda): \quad \begin{array}{lll} Z=2, & N=\nu, & P=\mu; \\ Z=1, & N=\nu+1, & P=\mu; \\ Z=1, & N=\nu, & P=\mu+1. \end{array}$$

If $A_{\tau+3}$ is non-singular (so that for $L_{\tau+2}(\lambda)$ the case $Z=2$ cannot arise), the outlay for $L_{\tau+3}(\lambda)$ is by Theorem V

$$(11) \qquad L_{\tau+3}(\lambda): \quad \begin{array}{lll} Z=0, & N=\nu+2, & P=\mu+1; \\ Z=0, & N=\nu+1; & P=\mu+2. \end{array}$$

That is, if $M_\tau M_{\tau+3}\neq 0$ while $M_{\tau+1}=M_{\tau+2}=0$, $L_{\tau+3}(\lambda)=0$ has two more or one more negative roots than $L_\tau(\lambda)=0$ according as M_τ and $M_{\tau+3}$ have the same sign or opposite signs.

Suppose, however, that $A_{\tau+1}$, $A_{\tau+2}$ and $A_{\tau+3}$ are singular while A_τ and $A_{\tau+4}$ are not. The possibilities for $L_{\tau+3}(\lambda)=0$ are then easily seen to be:

$$(12) \qquad L_{\tau+3}(\lambda): \quad \begin{array}{lll} Z=1, & N=\nu+1, & P=\mu+1; \\ Z=1, & N=\nu+2, & P=\mu; \\ Z=1, & N=\nu, & P=\mu+2; \end{array}$$

and for $L_{\tau+4}(\lambda)=0$:

$$(13) \qquad L_{\tau+4}(\lambda): \quad \begin{array}{lll} Z=0, & N=\nu+2, & P=\mu+2; \\ Z=0, & N=\nu+3, & P=\mu+1; \\ Z=0, & N=\nu+1, & P=\mu+3. \end{array}$$

If M_τ and $M_{\tau+4}$ are of the same sign, manifestly the first case in (13) is the only one that can arise, while if M_τ and $M_{\tau+4}$ are of opposite signs, either of the last two cases may arise and we cannot distinguish between them by the signs of the M's alone.

We therefore have the Theorem:

THEOREM VII. *If in the sequence* (9) $M_r\neq 0$ *and* $M_\tau M_{\tau+3}\neq 0$ *while* $M_{\tau+1}=M_{\tau+2}=0$, *then to the subsequence* M_τ, 0, 0, $M_{\tau+3}$ *we assign two variations and one permanence or one variation and two permanences of sign according as* M_τ *and* $M_{\tau+3}$ *have the same sign or opposite signs; and if* $M_\tau M_{\tau+4}\neq 0$ *while* $M_{\tau+1}=M_{\tau+2}=M_{\tau+3}=0$ *we assign to the subsequence* M_τ, 0, 0, 0, $M_{\tau+4}$ *exactly two variations and two permanences if* M_τ *and* $M_{\tau+4}$ *have the same sign, while in the contrary case the number of variations to be assigned may be either one or three.*

While the last theorem was proved only on the supposition that $M_\tau\neq 0$ for $\tau>0$ it is easy to verify that the results hold also for $\tau=0$.

The questions discussed in this section were studied originally by Frobenius,* and when two consecutive terms in the sequence (9) vanish the results that he arrived at by a very elaborate discussion are exactly the results that we have arrived at here. When three consecutive terms vanish and the adjacent M's have *opposite* signs, Frobenius points out that the signature of the matrix is not determined by the sequence (9) alone. But he does not seem to show that the signature *is* definitely determined when the adjacent M's have the *same* sign. More recently Franklin † attacked the same problem by a scheme similar to, but, it seems to the author, less explicit and less powerful than, the one used here, and he arrived at the same conclusions that Frobenius had previously arrived at. Still more recently and by an entirely different method the author ‡ obtained the results here given.

8. *A Sequence of Sturm Functions for the Equation* $L_n(\lambda)=0$. Let α and β be any two real numbers, neither a root of $L_n(\lambda)=0$. Since the characteristic roots of $A-\alpha I$ are less by α than the characteristic roots of A, it is clear that if ν_α is the number of characteristic roots $<\alpha$ of A, then ν_α is the number of negative roots of $A-\alpha I$. If in the sequence

$$1,\quad L_1(\alpha),\quad L_2(\alpha),\cdots,L_n(\alpha) \tag{14}$$

not more than two consecutive terms vanish (or if three consecutive terms vanish and the adjacent terms have the same sign), ν_α is equal to the number of variations of sign in the sequence, where if two or more consecutive terms vanish the number of variations is determined by theorem VII. Here a root of multiplicity m counts as m roots. Under the same restrictions if ν_β is the number of variations of sign in the sequence (14) with α replaced by β, then ν_β is the number of roots $<\beta$ of $L_n(\lambda)=0$. Hence for $\alpha<\beta$ $\nu_\beta-\nu_\alpha$ is the number of roots of $L_n(\lambda)=0$ between α and β. Without altering the number of variations of sign the order of the terms in (14) may be reversed thus furnishing a sequence in which the last one is always greater than zero. Such a sequence therefore

$$L_n(\lambda),\ L_{n-1}(\lambda),\cdots,L_1(\lambda),\ 1$$

may be thought of as constituting a sequence of Sturm functions § for the equation $L_n(\lambda)=0$.

THE UNIVERSITY OF NORTH CAROLINA.

* Frobenius, "Ueber das Trägsheitsgesetz der quadratischen Formen," *Crelle*, Vol. 114 (1895), pp. 198-199.

† Franklin, "A Theorem of Frobenius on Quadratic Forms," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 451-452.

‡ "On the Signature of a Quadratic Form," *Annals of Mathematics*, 2nd Series, Vol. 30 (1929), pp. 517-525.

§ Cf. Salmon, *Lessons on Higher Algebra*, Third Edition, Dublin (1876), p. 43.

Discontinuous Solutions in the Problem of Depreciation and Replacement.

By Henry H. Pixley.

1. *Introduction.* The mathematics of the problem of depreciation in economics has been the subject of recent papers by Hotelling * and by Roos.† Roos has developed a dynamical theory of depreciation and replacement and has formulated the problem of replacement for a single operating machine as a type of Lagrange problem in the calculus of variations. The expression which he maximizes is the sum of two definite integrals whose integrands are functions of variable end and corner values. He considers it as a single integral with an integrand which is discontinuous along a continuous curve of corners. The maximizing arc which he obtains is, however, continuous at the time of replacement. This means that the replacement machine starts at the time and at the rate of production at which the operating machine stops. In an actual case this is not necessarily true.

In this paper I develop a general theory corresponding to that of Roos without the assumptions of continuity at the time of replacement. In particular, an application is given in which the replacement machine is started at a time and rate different from those at which the operating machine stops.

2. *The replacement problem.* We consider a situation in which one machine operates from time t_1 to time w_1 at the rate of $u_1(t)$ units of output per unit time. Of the output of the machine $y_1(t)$ units are sold per unit time at a price $p_1(t)$ per unit. The total operating cost of the machine including depreciation is represented by the function $Q_1(u_1, u_1', p_1, p_1', t)$. A second machine operates from time $w_2 (\geqq w_1)$ to time t_2 with an output of $u_2(t)$ and a demand of $y_2(t)$ which sells at $p_2(t)$ per unit. The corresponding cost function is $Q_2(u_2, u_2', p_2, p_2', t)$.

Roos has shown that the total value, discounted to the time T, of the

* H. Hotelling, "A General Mathematical Theory of Depreciation," *Journal of American Statistical Association* (September, 1925).

† C. F. Roos, "A Mathematical Theory of Depreciation and Replacement," *American Journal of Mathematics*, Vol. 50 (January, 1928); Roos, "The Problem of Depreciation in the Calculus of Variations," *Bulletin of the American Mathematical Society*, Vol. 34 (1928), p. 218.

profits from a machine which operates from T_1 to T_2 plus the value at T of its scrap value at T_2 is

$$V(T)=\int_{T_1}^{T_2}[p(t)y(t)-Q(u,u',p,p',t)]\,E(T,t)\,dt+KE(T,T_2)$$

where K is the initial cost of the machine, and $E(T,t)$ is $\exp\left[-\int_T^t \delta(\nu)\,d\nu\right]$ in which $\delta(\nu)$ is the rate of increase of an invested sum s divided by s. The function $E(T,t)$ is a discount (or interest) factor which gives the value at T of the profits earned at t.* By this formula the total value at the time T of the two machines minus the value at T of the amounts necessary to replace the machines at w_2 and t_2 respectively is

$$\begin{aligned} &V_1(T)-K_1E(T,w_2)+V_2(T)-K_2E(T,t_2)\\ (1)\qquad &=\int_{t_1}^{w_1}(p_1y_1-Q_1)E(T,t)\,dt+K_1[E(T,w_1)-E(T,w_2)]\\ &\quad+\int_{w_1}^{w_2}(p_0y_0-Q_0)E(T,t)\,dt+\int_{w_2}^{t_2}(p_2y_2-Q_2)E(T,t)\,dt \end{aligned}$$

where the subscripts 1 and 2 denote functions of the operating and replacement machines respectively and the subscript 0 denotes functions for the period of replacement from w_1 to w_2. The cost function Q_0 represents any variable expense which occurs during the period of replacement and which may not be considered part of the constant K_2. The function Q_0 may be a function of w_1 and w_2.

For convenience we will drop the subscripts 1 and 2 for the present and let $u(t)$, $y(t)$, $p(t)$, and $Q(t)$ represent the rate of output, rate of demand, price, and cost of production, respectively, for the range $t_1 \leqq t \leqq t_2$. These functions may be discontinuous for the values $t=w_1, w_2$ but are continuous for all other values of t in the range $t_1 \leqq t \leqq t_2$. The functions $u(t)$, $y(t)$, and $p(t)$ are not in general independent, since y and p are related by an equation of demand, while y and u satisfy an equation of supply. If we assume that the demand equation is a first order differential equation of the form $\theta(y,y',p,p',t)=0$, and that the supply equation is $y=\xi(u,t)$, we can obtain by the elimination of y, a demand-supply equation †

$$(2)\qquad \phi(u,u',p,p',t)=0.$$

* Roos, "The Problem of Depreciation in the Calculus of Variations," *loc. cit.*, p. 221.

† Roos, "A Dynamical Theory of Economics," *The Journal of Political Economy*, Vol. 35 (October, 1927); Roos, "The Problem of Depreciation in the Calculus of Variations," *loc. cit.*, p. 222.

There will also in general be certain conditions which the end-points must satisfy which may be written in the form

$$(3) \qquad \psi_\mu\,[t_\sigma, u(t_\sigma), p(t_\sigma), w_\sigma, u(w_\sigma), p(w_\sigma \pm 0)] = 0, \qquad (\mu = 1, \cdots, p \leqq 14),$$

where σ may take both of the values 1, 2 in each of the p equations.

We will eliminate y from the function (1) by means of the equation $y = \xi(u, t)$. Then if we assume that this function is to be maximized, our problem is that of finding among the arcs $u(t)$, $p(t)$, satisfying the equation (2), and whose end-points satisfy equations (3), a set which maximizes this expression (1).

3. *The general problem.* We will now state a more general problem of which the problem of the preceding paragraph is a special case. We will need to consider a class of arcs, $y_i = y_i(x)$, $(i = 1, \cdots, n)$, which are defined for $x = s_1$, $x = s_2$, and $x = s_3$, where $x_1 \leqq s_1 \leqq x_2$, $x_3 \leqq s_2 \leqq x_4$, $x_5 \leqq s_3 \leqq x_6$. We will represent these three intervals by the letters X_1, X_2, and X_3, respectively. Our general problem is that of *finding among those discontinuous arcs, $y_i = y_i(x)$, $(i = 1, \cdots, n)$, of the above class which satisfy certain differential equations*

$$(4) \qquad \phi_\alpha(x, y, y') = 0, \qquad (\alpha = 1, \cdots, m < n),$$

for all x in X_1, X_2, and X_3, and whose points at the ends of these intervals satisfy the end equations

$$(5) \quad \psi_\mu\,[x_\rho, y(x_\rho)] = 0, \qquad (\mu = 1, \cdots, p \leqq 6n + 6;\ \rho = 1, \cdots, 6),$$

one which maximizes an expression

$$(6) \quad I = \int_{x_1}^{x_2} f\,[s_1, y, y', x_\rho, y(x_\rho)]\,ds_1 + \int_{x_3}^{x_4} g\,[s_2, y, y', x_\rho, (x_\rho)]\,ds_2 + \int_{x_5}^{x_6} h\,[s_3, y, y', x_\rho, y(x_\rho)]\,ds_3.$$

where (y, y') represents the set $(y_1, \cdots, y_n, y_1', \cdots, y_n')$, $[x_\rho, y(x_\rho)]$ represents the set

$[x_1, y_1(x_1), \cdots, y_n(x_1), x_2, y_1(x_2), \cdots, y_n(x_2), \cdots, x_6, y_1(x_6), \cdots, y_n(x_6)]$,
and primes denote differentiation with respect to x.

We assume that:

(a) the functions $y_i(x)$ defining the maximizing arc E are continuous

in each of the intervals X_1, X_2, X_3, and have continuous derivatives in these intervals except at a finite number of values of x;

(b) in a neighborhood R of the values (x, y, y') on the arc E the functions f, g, h, and ϕ_α have continuous derivatives up to and including those of the second order;

(c) at every element (x, y, y') on E the $m \times n$-dimensional matrix $\|\phi_{\alpha y_i'}\|$ has rank m;

(d) the functions ψ_μ have continuous derivatives up to and including those of the second order near the end values $[x_\rho, y(x_\rho)]$ and at these values the $p \times (6n + 6)$-dimensional matrix

$$\| \psi_{\mu x_\rho} \quad \psi_{\mu y_\rho} \| \tag{7}$$

has rank p, where $y_\rho = y(x_\rho)$, $(\rho = 1, \cdots, 6)$, and the subscripts y_i', x_ρ, y_ρ denote partial derivatives.*, †

For the general problem as here stated certain necessary conditions for a solution can be obtained by methods which are essentially those given by Bliss for the problem with a continuous integrand,* and which have been extended by Roos to the case of a discontinuous integrand.† In each of these treatments the solution is sought in a class of continuous arcs.

4. *Admissible arcs and variations.* An arc $y_i = y_i(x)$, $(i = 1, \cdots, n)$, defined over the intervals X_1, X_2, X_3 will be called an *admissible arc* if it has the continuity properties (a); if all of its elements (x, y, y') lie in R, and if it satisfies the differential equations (4).

If a one-parameter family of admissible arcs

$$y_i = y_i(x, b), \tag{8}$$

$[i = 1, \cdots, n;\ x_1(b) \leqq x \leqq x_2(b);\ x_3(b) \leqq x \leqq x_4(b);\ x_5(b) \leqq x \leqq x_6(b)]$ containing a particular admissible arc E for the parameter value $b = 0$ be given, the functions $\eta_i(x) = \partial y_i(x, 0)/\partial b$, $\xi_\rho = \partial x_\rho(0)/\partial b$ are called *variations of the family along E.*

The equations of variation on the arc E for the functions ϕ_α are defined by

$$\Phi_\alpha(\eta, \eta') = \phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i' = 0, \qquad (\alpha = 1, \cdots, m), \tag{9}$$

* G. A. Bliss, "Lectures on the Problem of Lagrange in the Calculus of Variations," *University of Chicago* (1925), mimeographed by O. E. Brown, University of Chicago.

† Roos, "General Problem of Minimizing an Integral with Discontinuous Integrand," *Transactions of the American Mathematical Society*, Vol. 31, (January, 1929), (hereafter referred to as "General Problem").

where the coefficients ϕ_{ay_i}, $\phi_{ay_i'}$ have as arguments the functions $y_i(x)$ defining the arc E and the functions η_i, η_i' are, of course, defined only for values of x in the intervals X_1, X_2, X_3.

Similarly we define the equations of variation on the arc E for the functions ψ_μ to be

$$\Psi_\mu(\xi, \eta) = \psi_{\mu x_\rho}\xi_\rho + \psi_{\mu y_{i\rho}} dy_i[x_\rho(0), 0]/db, \tag{10}$$

where in equations (9) and (10) i is an umbral index with range $1, \cdots, n$, and ρ is umbral with range $1, \cdots, 6$, according to the convention that whenever a subscript appears twice in a term that term is to be summed for all values of the subscript. The functions Ψ_μ are clearly functions of ξ_ρ and η_i since

$$dy_i[x_\rho(0), 0]/db = y_{i\rho}'\xi_\rho + \eta_i(x_\rho), \quad (\rho = 1, \cdots, 6, \text{ not umbral}). \tag{11}$$

A set of arbitrary constants ξ_ρ and functions $\eta_i(x)$ with the continuity properties described in (a) and satisfying the equations of variation (9) will be called a *set of admissible variations,* a definition which we will find useful since

*For every set of admissible variations ξ_ρ, $\eta_i(x)$ along the arc E there exists a one-parameter family (8) of admissible arcs containing E for the value $b = 0$ and having the set ξ_ρ, $\eta_i(x)$ as its variations along E. For this family the functions $y_i(x, b)$ are continuous on each of the intervals X_1, X_2, X_3 and have continuous derivatives with respect to b for all values (x, b) near those defining E, and the derivatives $dy_i(x, b)/dx$ have the same property except, possibly, at the values of x defining corners of E.**

5. *First necessary conditions.* If we substitute the one-parameter family of admissible arcs, (8), containing E for $b = 0$, in the expression I, differentiate I with respect to b, and set $b = 0$, we obtain the first variation of I along the arc E

$$\begin{aligned} I_1(\xi, \eta) = & \int_{x_1}^{x_2} (f_{y_i}\eta_i + f_{y_i'}\eta_i')\, ds_1 \\ & + \int_{x_3}^{x_4} (g_{y_i}\eta_i + g_{y_i'}\eta_i')\, ds_2 + \int_{x_5}^{x_6} (h_{y_i}\eta_i + h_{y_i'}\eta_i')\, ds_3 \\ & + K_{i\rho}(f, g, h)\, dy_i[x_\rho(0), 0]/db + L_\rho(f, g, h)\xi_\rho, \end{aligned} \tag{12}$$

* For proof see Roos, "General Problem," *loc. cit.*, p. 61. See also Bliss, "Lectures, etc.," *loc. cit.*, p. 4. The theorem stated above is an obvious extension of the one proved by Roos.

where $K_{i\rho}(f,g,h)=\int_{x_1}^{x_2} f_{y_i\rho}ds_1+\int_{x_3}^{x_4} g_{y_i\rho}ds_2+\int_{x_5}^{x_6} h_{y_i\rho}ds_3$;

$$L_\rho(f,g,h)=f_\rho+\int_{x_1}^{x_2} f_{x\rho}ds_1+\int_{x_3}^{x_4} g_{x\rho}ds_2+\int_{x_5}^{x_6} h_{x\rho}ds_3;\,*$$

$f_1=-f(x_1)$, $f_2=f(x_2)$, $f_3=-g(x_3)$, $f_4=g(x_4)$, $f_5=-h(x_5)$, $f_6=h(x_6)$; $f(x_1)$ is the value of the function at the end-point of the arc E corresponding to $x=x_1$ and the other functions, f_ρ, are similarly defined; i is an umbral index with range $1,\cdots,n$, and ρ is umbral with range $1,\cdots,6$; and the subscripts y_i, y_i', $y_{i\rho}$, x_ρ denote partial derivatives.

Following the methods of Bliss and Roos it can be proved by means of this first variation that: *For every maximizing arc for the above problem there exist sets of constants c_{i1}, c_{i2}, c_{i3}, $(i=1,\cdots,n)$, and functions*

$$F(s_1,y,y',x_\rho,y_{i\rho},\lambda_0,\lambda_\alpha)=\lambda_0 f+\lambda_\alpha\phi_\alpha,$$
$$G(s_2,y,y',x_\rho,y_{i\rho},\lambda_0\ \lambda_\alpha)=\lambda_0 g+\lambda_\alpha\phi_\alpha,$$
$$H(s_3,y,y',x_\rho,y_{i\rho},\lambda_0\ \lambda_\alpha)=\lambda_0 h+\lambda_\alpha\phi_\alpha,\qquad(\alpha=1,\cdots,m;\ \text{umbral}),$$

such that the equations

$$(13)\quad F_{y_i'}=\int_{x_1}^{s_1}F_{y_i}ds_1+c_{i1},\ G_{y_i'}=\int_{x_3}^{s_2}G_{y_i}ds_2+c_{i2},\ H_{y_i'}=\int_{x_5}^{s_3}H_{y_i}ds_3+c_{i3},$$

are satisfied at every point of E. The constant λ_0 and the functions $\lambda_\alpha(x)$, $(\alpha=1,\cdots,m)$, are not all identically zero on the intervals X_1, X_2, X_3, and are continuous except possibly at values of x defining corners of E. Furthermore, the end values of E must be such that all determinants of order $p+1$ of the matrix

$$(14)\qquad \begin{Vmatrix} N_\rho(x_\nu) & M_{i\rho}(x_\nu) \\ \psi_{\mu x_\rho} & \psi_{\mu y_{i\rho}} \end{Vmatrix}$$

vanish, where

$$M_{i\rho}(x_\nu)=F_{i\rho}+K_{i\rho}(F,G,H),\quad N_\rho(x_\nu)=-F_{i\rho}y'_{i\rho}+L_\rho(F,G,H);$$

$F_{i1}=-F_{y_i'}(x_1)$, $F_{i2}=F_{y_i'}(x_2)$, $F_{i3}=-G_{y_i'}(x_3)$, $F_{i4}=G_{y_i'}(x_4)$, $F_{i5}=-H_{y_i'}(x_5)$, $F_{i6}=H_{y_i'}(x_6)$; $F_{y_i'}(x_1)$ denotes a derivative with respect to y_i' evaluated at the end-point of E defined by $x=x_1$, and the other functions $F_{i\rho}$ are similarly defined; i is umbral with the range $1,\cdots,n$; $\rho=1,\cdots,6$ and ρ is not umbral; and (x_ν) denotes the set $(x_1,\cdots,x_6)$.

---

* The notation here is suggested by Roos. See "General Problem," *loc. cit.*, p. 62.

6. *The maximizing arcs for the replacement problem.* In the problem stated in § 2 we will assume that the relation, $y = \xi(u, t)$, between the rate of demand and the rate of supply is of the form $y(t) = \alpha_\sigma u(t) + \beta_\sigma(t)$, and furthermore that the demand is a linear function of the price and the rate of change of price, $y(t) = d_\sigma p(t) + e_\sigma(t) + k_\sigma p'(t)$. Then the demand-supply equation, (2), becomes

$$\phi_{1\sigma} = u - a_\sigma p(t) - b_\sigma(t) - h_\sigma p'(t) = 0, \qquad (\sigma = 1, 2), \tag{15}$$

where, as in § 2, $\sigma = 1$, and $\sigma = 2$ denote functions of the operating and replacement machines, respectively; $a_\sigma = d_\sigma/\alpha_\sigma$, $b_\sigma = (e_\sigma + \beta_\sigma)/\alpha_\sigma$, $h_\sigma = k_\sigma/\alpha_\sigma$; and it must be remembered that the forms of the expressions represented by $u(t)$, $y(t)$, and $p(t)$ are not in general the same for the operating and replacement machines. For the period of replacement from the time w_1 to time w_2 we have $u \equiv 0$, and in the place of equation (15) we use the demand equation $y(t) = d_0 p(t) + e_0(t) + k_0 p'(t)$. If in addition we know the initial time, $t_1 = T_1$, the rate and price of output at time t_1, the rate of output at time w_2, and the time which elapses between w_1 and w_2, the conditions (3) are

$$\begin{gathered}\psi_1 = t_1 - T_1 = 0, \quad \psi_2 = u(T_1) - U_1 = 0, \quad \psi_3 = p(T_1) - P_1 = 0, \\ \psi_4 = w_1 - w_2 + W = 0, \quad \psi_5 = u(w_2) - U_2 = 0,\end{gathered} \tag{16}$$

in which T_1, U_1, P_1, W, U_2 are known constants.

Let us also suppose that the cost function Q is expressible by means of the forms

$$\begin{aligned} Q_\sigma(u, u', p, p', t) &= A_\sigma u^2 + B_\sigma u + C_\sigma + D_\sigma u'^2 + E_\sigma p'^2 + F_\sigma u' + G_\sigma p' + H_\sigma p^2 + I_\sigma p, \\ &\qquad\qquad\qquad\qquad (\sigma = 1, 2), \\ Q_0(p, p', t) &= \qquad C_0(t) \qquad + E_0 p'^2 \qquad + G_0 p' + H_0 p^2 + I_0 p. \end{aligned}$$

The parameters a_σ, b_σ, α_σ, β_σ, d_0, e_0, k_0, h_σ, $A_\sigma, \cdots, I_\sigma$, C_0, E_0, G_0, H_0, I_0 are either known functions of the time or constants. In the following solution we will consider all of them except b_σ, C_0, and e_0 as constants for simplicity of the solution, although the problem can be solved when they are functions of t. We will also consider $\delta(\nu)$ a constant.

Our problem may now be stated as that of: *Finding among the arcs* $u(t)$, $p(t)$ *satisfying a demand-supply equation* (15), *and whose end-points satisfy equations* (16), *a set which maximizes the expression* (1), *which may be written*

$$I = \int_{t_1}^{w_1} (\alpha_1 p_1 u_1 + \beta_1 p_1 - Q_1) E(T, t)\, dt$$

$$(17) \quad + \int_{w_1}^{w_2} \left[\frac{K_1[E(T, w_1) - E(T, w_2)]}{(w_2 - w_1)} + \{(d_0 p_0 + e_0 + k_0 p_0') p_0 - Q_0\} E(T, t) \right] dt$$

$$+ \int_{w_2}^{t_2} (\alpha_2 p_2 u_2 + \beta_2 p_2 - Q_2) E(T, t)\, dt.$$

This is a special case of the general problem stated in § 3 where $x_1 = t_1$, $w_1 = x_2 = x_3$, $x_4 = x_5 = w_2$, $x_6 = t_2$, $y_1(x) = u(t)$, $y_2(x) = p(t)$, ϕ_1 is ϕ_{11} for $t_1 \leqq t \leqq w_1$, u for $w_1 \leqq t \leqq w_2$, and ϕ_{12} for $w_2 \leqq t \leqq t_2$, and f, g, and h correspond to the three integrand functions. Therefore the arcs $u(t)$, $p(t)$ with their end-points must satisfy the equations (13) and the tranversality conditions (14).

If we define F, G, and H by the equations*

$$\begin{aligned} F &= [\alpha_1 pu + \beta_1 p - A_1 u^2 - B_1 u - C_1 - D_1 u'^2 - E_1 p'^2 - F_1 u' \\ &\qquad - G_1 p' - H_1 p^2 - I_1 p + \lambda_{11}(u - a_1 p - h_1 p' - b_1(t))]\, E(T, t) \\ G &= K_1[E(T, w_1) - E(T, w_2)]/(w_2 - w_1) \\ &\qquad + [(d_0 p + e_0 + kp')p - C_0(t) - E_0 p'^2 - G_0 p' - H_0 p^2 - I_0 p] E(T, t) \\ H &= [\alpha_2 pu + \beta_2 p - A_2 u^2 - B_2 u - C_2 - D_2 u'^2 - E_2 p'^2 - F_2 u' \\ &\qquad - G_2 p' - H_2 p^2 - I_2 p_2 + \lambda_{12}(u - a_2 p - h_2 p' - b_2(t))]\, E(T, t) \end{aligned}$$

we obtain

$$\begin{aligned} \partial F/\partial u &= (\alpha_1 p - 2A_1 u - B_1 + \lambda_{11}) E(T, t), \\ \partial F/\partial u' &= (-2D_1 u' - F_1) E(T, t), \\ \partial F/\partial p &= (\alpha_1 u + \beta_1 - 2H_1 p - I_1 - a_1 \lambda_{11}) E(T, t), \\ \partial F/\partial p' &= (-2E_1 p' - G_1 - h_1 \lambda_{11}) E(T, t). \end{aligned}$$

The Euler-Lagrange equations in their classical form $dF_{y_i'}/dx = F_{y_i}$ are obtained from equations (13) by differentiation, and in our case these conditions are

$$(18) \quad \begin{aligned} &-2D_1 u'' + (2D_1 u' + F_1)\delta = \alpha_1 p - 2A_1 u - B_1 + \lambda_{11} \\ &-2E_1 p'' - h_1 \lambda_{11}' + (2E_1 p' + G_1 + h_1 \lambda_{11})\delta = \alpha_1 u + \beta_1 - 2H_1 p - I_1 - a_1 \lambda_{11} \end{aligned}$$

from each of which the common factor $E(T, t)$ has been removed. Solving the first of these equations for λ_{11} and substituting its value in the second, we obtain

$$\begin{aligned} &2h_1 D_1 u''' - 2(a_1 + 2h_1\delta) D_1 u'' + 2(-h_1 A_1 + \gamma_1 \delta D_1) u' + (2\gamma_1 A_1 - \alpha_1) u - 2E_1 p'' \\ &\qquad + (\alpha_1 h_1 + 2\delta E_1) p' + (-\alpha_1 \gamma_1 + 2H_1) p + \gamma_1 (B_1 + \delta F_1) + \delta G_1 + I_1 - \beta_1 = 0 \end{aligned}$$

where $\gamma_1 = a_1 + h_1 \delta$.

---

* See Roos, "A Mathematical Theory of Depreciation and Replacement," *loc. cit.*, p. 153.

Replacing u and its derivatives in this equation by their values in terms of p obtained from the demand-supply equation (15), we have the differential equation

$$(19)\quad L_{14}D_t^4p + L_{13}D_t^3p + L_{12}D_t^2p + L_{11}D_tp + L_{10}p + p_{10} = L_1(b_1, b_1', b_1'', b_1''')$$

in which

$$L_{14} = 2h_1^2D_1,\ L_{13} = -4h_1^2\delta D_1,\ L_{12} = -2h_1^2A_1 + 2(-a_1\gamma_1 + h_1^2\delta^2)D_1 - 2E_1,$$
$$L_{11} = 2\delta(h_1^2A_1 + a_1\gamma_1D_1 + E_1),\quad L_{10} = 2a_1\gamma_1A_1 + 2H_1 - 2a_1\alpha_1 - h_1\alpha_1\delta,$$
$$L_1(b_1, b_1', b_1'', b_1''') = -2h_1D_1b_1''' + 2(a_1 + 2h_1\delta)D_1b_1'' + 2(h_1A_1 - \gamma_1\delta D_1)b_1' + (-2\gamma_1A_1 + \alpha_1)b_1,$$
$$p_{10} = \gamma_1(B_1 + \delta F_1) + \delta G_1 + I_1 - \beta_1.$$

Since this is a linear differential equation with constant coefficients its solution depends upon the roots, m_{11}, m_{12}, m_{13}, m_{14}, of the algebraic equation $L_{14}m^4 + L_{13}m^3 + L_{12}m^2 + L_{11}m + L_{10} = 0$. If these roots are all distinct and if $\bar{p}_1(t)$ is a particular solution of equation (19), then its general solution is

$$(20)\qquad p_1 = \bar{p}_1(t) + K_{11}e^{m_{11}t} + K_{12}e^{m_{12}t} + K_{13}e^{m_{13}t} + K_{14}e^{m_{14}t},$$

where the constants K_{11}, K_{12}, K_{13}, K_{14} are arbitrary.

The determination of $\bar{p}_1$ depends, of course, on $b_1(t)$. An interesting form of $b_1(t)$ is the general solution of the homogeneous linear differential equation $L_1(b_1, b_1', b_1'', b_1''') = 0$. The auxiliary equation in this case is $-2h_1D_1\mu^3 + 2(a_1 + 2h_1\delta)D_1\mu^2 + 2(h_1A_1 - \gamma_1\delta D_1)\mu - 2\gamma_1A_1 + \alpha_1 = 0$ and if its roots μ_{11}, μ_{12}, μ_{13} are all distinct the general solution of the equation $L_1 = 0$ is

$$(21)\qquad b_1(t) = \bar{K}_{11}e^{\mu_{11}t} + \bar{K}_{12}e^{\mu_{12}t} + \bar{K}_{13}e^{\mu_{13}t},$$

where $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ are arbitrary constants. The constants $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ are at the disposal of the operator in forming a satisfactory demand-supply equation (15). Hence for this form of $b_1(t)$ our demand-supply equation has five arbitrary constants and at the same time gives us a solution for $p_1(t)$ which can always be expressed explicitly in the form (20). Since $L_1 = 0$, the particular solution may be taken $\bar{p}_1 = -p_{10}/L_{10}$.

It may appear that the price p_1 as given in (20) is independent of the constants $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ in $b_1(t)$. However, in practice, for any change in $\bar{K}_{11}$, $\bar{K}_{12}$, $\bar{K}_{13}$ one would probably choose different values for a_1 and h_1 in the demand-supply equation (15), and p_1 is a function of these constants.

The differential equation which gives $p_2(t)$ is formally like (20), its coefficients being functions of a_2, $b_2(t)$, h_2, α_2, β_2, $A_2, \cdots, I_2$. If $b_2(t)$ is defined by an equation similar to (21), then

$$p_2 = \bar{p}_2(t) + K_{21}e^{m_{21}t} + K_{22}e^{m_{22}t} + K_{23}e^{m_{23}t} + K_{24}e^{m_{24}t}$$

in which the K's and m's have meanings analogous to those in equation (20). As soon as p_1 and p_2 are known we have u_1 and u_2 from the demand-supply equation (15).

The differential equation which gives $p_0(t)$ is $dG_{p'}/dt = G_p$, which in terms of the coefficients of the cost function becomes

$$-2E_0p'' + 2\delta E_0p' + (2H_0 - 2d_0 - k_0\delta)p + \delta G_0 + I_0 = e_0.$$

If m_{01} and m_{02} are the roots of $2E_0m^2 - 2\delta E_0m - 2H_0 + 2d_0 + k_0\delta = 0$, the solution for p_0 may be written

$$p_0 = \bar{p}_0(t) + K_{01}e^{m_{01}t} + K_{02}e^{m_{02}t},$$

where $\bar{p}_0(t)$ is any solution of the differential equation and K_{01}, K_{02} are arbitrary constants. In particular, if e_0 is a constant, this solution may be taken $\bar{p}_0 = (e_0 - G_0\delta - I_0)/2H_0$. However, the finding of a particular solution does not depend on e_0 being a constant since there are many functions of t which put in the place of e_0 would yield a particular solution easily.

7. *Determining the end values.* We now use the conditions on the end values t_1, $u_1(t_1)$, $p_1(t_1)$, w_1, $u_1(w_1)$, $p_1(w_1)$, $p_0(w_1)$, w_2, $p_0(w_2)$, $u_2(w_2)$, $p_2(w_2)$, t_2, $u_2(t_2)$, $p_2(t_2)$, to determine the constants t_σ, w_σ, $K_{\sigma1}$, $K_{\sigma2}$, $K_{\sigma3}$, $K_{\sigma4}$, K_{01}, K_{02}. These end values are subject to the transversality conditions (14). Since $w_1 = x_2 = x_3$, $x_4 = x_5 = w_2$, we must add the equations $\psi_6 = x_2 - x_3 = 0$, $\psi_7 = x_4 - x_5 = 0$, to the known end conditions (16) in evaluating the matrix (14). We now find that every determinant of order 8 of the $(8+18)$-dimensional matrix

$$\left\|\begin{array}{ccccccccc} N_1 & -F_{u_1'}(t_1) & -F_{p_1'}(t_1) & N_2 & N_3 & N_4 & N_5 & -H_{u_2'}(w_2) & c_k \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{array}\right\|$$

$(k=1,\cdots,10)$, must vanish. In this matrix $c_1=F_{u_1'}(w_1)$, $c_2=F_{p_1'}(w_1)$, $c_3=-G_{u_0'}(w_1)\equiv 0$, $c_4=-G_{p_0'}(w_1)$, $c_5=G_{u_0'}(w_2)\equiv 0$, $c_6=G_{p_0'}(w_2)$, $c_7=-H_{p_2'}(w_2)$, $c_8=N_6=-H_{u_2'}(t_2)u_2'(t_2)-H_{p_2'}(t_2)p_2'(t_2)+H(t_2)$, $c_9=H_{u_2'}(t_2)$, $c_{10}=H_{p_2'}(t_2)$. If it is assumed that the time T to which all profits are discounted is t_1, necessary and sufficient conditions that every determinant of order 8 vanish are

$$N_2(w_1,w_2)+N_3(w_1)+N_4(w_2)+N_5(w_1,w_2)=0,\quad c_k=0,$$

$(k=1,\cdots,10)$, which are equivalent to the following equations:

$$\begin{aligned}N_2+N_3+N_4+N_5&+c_1u_1'(w_1)+c_2p_1'(w_1)+c_4p_0'(w_1)+c_6p_0'(w_2)+c_7p_2'(w_2)\\ &=F(w_1)-G(w_1)+G(w_2)-H(w_2)+H_{u_2'}(w_2)u_2'(w_2)+\int_{w_1}^{w_2}(G_{w_1}+G_{w_2})\,dt=0,\end{aligned}$$

$$\begin{aligned}&c_1=[-2D_1u_1'(w_1)-F_1]\,E(t_1,w_1)=0,\\ &c_2=[-2E_1p_1'(w_1)-G_1-h_1\lambda_{11}(w_1)]\,E(t_1,w_1)=0,\\ &c_3\equiv c_5\equiv 0,\\ (22)\qquad &c_4=[2E_0p_0'(w_1)+G_0]\,E(t_1,w_1)=0,\\ &c_6=[-2E_0p_0'(w_2)-G_0]\,E(t_1,w_2)=0,\\ &c_7=[2E_2p_2'(w_2)+G_2+h_2\lambda_{12}(w_2)]\,E(t_1,w_2)=0,\\ &c_8+c_9u_2'(t_2)+c_{10}p_2'(t_2)=H(t_2)=0,\\ &c_9=[-2D_2u_2'(t_2)-F_2]\,E(t_1,t_2)=0,\\ &c_{10}=[-2E_2p_2'(t_2)-G_2-h_2\lambda_{12}(t_2)]\,E(t_1,t_2)=0.\end{aligned}$$

Since u and p are expressible in exponentials in t, the integration indicated in the first of these equations can be performed without difficulty and the explicit expression in terms of the given constants can then be exhibited as has been done in the other equations.

The fourteen constants t_σ, w_σ, $K_{\sigma1}$, $K_{\sigma2}$, $K_{\sigma3}$, $K_{\sigma4}$, K_{01}, K_{02}, can now be determined by the five equations (16) and the nine equations (22), the system (22) giving us only nine equations since c_3 and c_5 are identically zero.

Interesting interpretations can be given some of the end-conditions. Since $H(t)$ represents the profits per unit time from the replacement machine, the condition $H(t_2)=0$ means that the replacement machine should be run until the amount of money received for the goods sold equals the cost of production at that time. From the condition $u_1'(w_1)=-F_1/2D_1$ the slope of the production curve at the time the operating machine is scrapped is seen to be a constant which depends only on the coefficients of the cost function of the machine. Roos has shown that in typical cases we have $D_1>0$ and $F_1\geqq 0$.*

* See Roos, "Some Problems of Business Forecasting," *Proceedings of the National Academy of Sciences*, Vol. 15 (March, 1929), p. 190.

Hence the rate of production is decreasing at this time. Similar conditions on the rate of production at the time of scrapping the replacement machine follow from the equation $u_2'(t_2) = -F_2/2D_2$.

8. *Other forms of the problem.* Evans has suggested that in certain cases the demand depends partly on the seasons, and he has given a form of the demand-supply equation which involves a periodic term as follows: $y = ap + b + b' \cos kt + h dp/dt$, $b' < b$, a, b, b', h all constants.* It will be noticed that if in equation (21), $\mu_{11} = 0$, and μ_{12}, μ_{13} are pure imaginaries (hence, equal except for sign, since we assume all coefficients to be real) the resulting demand-supply equation is in Evans' form.

There is also a variety of other forms of this function $b(t)$ in the demand-supply equation which give a readily integrable differential equation (19). In particular, if b is any exponential in the first power of t, or a polynomial in t, or a constant is this true.

The end equations (16) could also be replaced by other conditions without altering the analysis of the problem. It will be noticed that the conditions (22) can be simplified by assuming that more of the end values are known constants.

* G. C. Evans, "The Mathematical Theory of Economics," *American Mathematical Monthly*, Vol. 32 (1925), p. 108.

A Prepared System for Two Quinary Quadratic Forms.

By J. WILLIAMSON.

Introduction. In a previous paper,* a prepared system was determined, in terms of which every concomitant of two quadratics in n variables could be expressed, if the concomitants were multiplied by suitable invariant factors. In this paper we determine a prepared system, for the case $n=5$, in terms of which every concomitant can be expressed, without being multiplied by an invariant factor. We find that eight new factors must be added to the $2^5-1=31$ factors already determined, giving a total of 39. In addition a complete list of several types of irreducible concomitants is obtained.

We use the notation of the previous paper throughout except that, for convenience in printing, primes are now used to denote determinantal permutations; i. e. $(ab'c')d_x'$ is used instead of $(a\dot{b}\dot{c})\dot{d}_x$ to denote the series $(abc)d_x-(abd)c_x-(adc)b_x$. Furthermore, for the five sets of cogredient point variables, that are necessary for this discussion, we use x, y, z t, w, and write P, p, and u for the compound coördinates π_2, π_3, and π_4 respectively.†

In the first two sections the results are listed, while the remaining sections are devoted to their determination.

1. *The Prepared System.* This system consists of 5 x-factors, 5 u-factors, 10 P-factors, 10 p-factors, the factor (12345), 3 px-factors, 3 Pu-factors and 2 xu-factors. A complete list of these factors is given below.

$1_x=a_x,\quad 5_x,\quad 2_x=(A\rho x)=a_\rho' b_x',\quad 4_x,\quad 3_x=(A_3R_3x)=(a'b'R_3)c_x'$

$(12)=a_\rho(AP),\quad (54),\quad (13)=(aA_3R_3P)=(a'b'R_3)(ac'P),\quad (53),$

$(14)=(aR\alpha P)=r_\alpha'(as'P),\quad (52),\quad (15)=(arP),$

$(23)=(AR_3)(A_3\rho P)=(AR_3)a_\rho'(b'c'P),\quad (43),$

$(24)=(A\rho R\alpha P)=a_\rho'\overline{r}_\alpha(b'\overline{s}P);$

$(123)=a_\rho(AR_3)(A_3p),\quad (543),\quad (124)=a_\rho(AR\alpha p)=a_\rho r_\alpha'(As'p),$

$(542),\quad (125)=a_\rho(Arp),\quad (541),\quad (143)=(RA_3)(aR_3\alpha p)=(RA_3)r_\alpha'(as't'p),$

$(523),\quad (135)=(aA_3R_3rp)=(a'b'R_3)(ac'rp),$

$(234)=(AR_3)(RA_3)(\alpha\rho p)=(AR_3)(RA_3)a_\rho'(b'c'd'p);$

* J. Williamson, "A Special Prepared System for Two Quadratics in n variables," *American Journal of Mathematics*, Vol. 52 (April, 1930), pp. 399-412.

† *Loc. cit.*, §§ 1 and 2.

$(1234) = a_\rho (AR_3)(RA_3)u_\alpha,\quad (5432),\quad (1235) = a_\rho (AR_3)(A_3 ru),$
$(5431),\quad (1254) = a_\rho r_\alpha (ARu);\quad (12345) = a_\rho r_\alpha (AR_3)(RA_3);$
$(12, 54) = 1_x'(2'54) = (125')4_x' = a_\rho r_\alpha (ARpx) a_\rho r_\alpha a_x'(b'Rp) = a_\rho r_\alpha s_x'(Ar'p),$
$(12, 43) = 1_x'(2'43) = (124')3_x' = a_\rho (RA_3)(AR_3 \alpha px) = a_\rho (RA_3) r_x' s_\alpha'(At'p),$
$(54, 23) = 5_x'(4'23) = (542')3_x';$
$(123, 543) = (2'3)(1'543) = (4'3)(1235') = a_\rho (AR_3) r_\alpha (RA_3)(A_3 R_3 Pu),$
$= a_\rho (AR_3) r_\alpha (RA_3)(b'c'P)(a'R_3 u) = a_\rho (AR_3) r_\alpha (RA_3)(r's'P)(A_3 t'u),$
$(123, 154) = (12')(1543') = (14')(1235') = a_\rho (AR_3) r_\alpha (A_3 aRPu),$
$= a_\rho (AR_3) r_\alpha (a'b'P)(aRc'u) = a_\rho (AR_3) r_\alpha (as'P)(A_3 r'u),$
$(543, 512) = (54')(5123') = (52')(5431');$
$(12, 543) = 1_x'(5432') = 3_x'(5'4'12) = a_\rho r_\alpha (RA_3)(R_3 Aux),$
$= a_\rho r_\alpha (RA_3) a_x'(R_3 b'u) = a_\rho r_\alpha (RA_3) t_x'(r's'Au),$
$(54, 123) = 5_x'(1234') = 3_x'(1'2'54).$

In the above list A, R, α, ρ have been written for A_2, R_2, A_4, R_4 respectively and $A = ab$, $A_3 = abc$, $R = rs$, $R_3 = rst$. When two factors are similar, only one has been defined, since the other may be obtained by replacing a, A, A_3, α by r, R, R_3, ρ respectively.

2. *Complete list of irreducible concomitants of several types.*

6 invariants: $(a\alpha)^2$, $(a\rho)^2$, $(AR_3)^2$, $(RA_3)^2$, $(r\alpha)^2$, $(r\rho)^2$.

6 covariants: 5 quadratics i_x^2 and 1 quintic $(12345)1_x 2_x 3_x 4_x 5_x$.

6 contravariants: 5 quadratics $(ijkm)^2$,
1 quintic $(1234)(1235)(1245)(1345)(2345)$.

20 complexes containing the variable P:
10 quadratics $(ij)^2$, 10 cubics $(ij)(jk)(ki)$.

20 complexes containing the variable p:
10 quadratics $(ijk)^2$, 10 cubics $(12345)(ijk)(ijm)(ijn)$.

44 mixed forms containing u and x:
5 of orders 1 in u and 1 in x, $(12345)(1234)5_x S$,
$(12345)(1245)3_x$, $(12345)(54, 123)S$.
5 of orders 1 in u and 4 in x, $(ijkm)i_x j_x k_x m_x$.
5 of orders 4 in u and 1 in x,
$(12345)(mijk)(mijn)(mikn)(mjkn)m_x$.
10 of orders 2 in u and 3 in x, $(12345)(ijkm)(ijkn)i_x j_x k_x$.
10 of orders 3 in u and 2 in x, $(mnkj)(mnij)(mnik)m_x n_x$.
2 of orders 3 in u and 3 in x, $(12345)(1245)(1345)(2345)1_x 2_x 3_x S$,
2 of orders 3 in u and 4 in x, $(12, 543)(1245)(1345)1_x 4_x 5_x S$.

4 of orders 4 in u and 4 in x,
$(12, 543)(12345)(1234)(1245)(1345)1_x 4_x S$,
$(12, 543)(12345)(1235)(1245)(1345)1_x 5_x S$,
1 of orders 5 in u and 4 in x,
$(12, 543)(54, 123)(1235)(1245)(1345)1_x 5_x$.

67 mixed forms containing P and x:

10 of orders 1 in P and 2 in x, $(ij)i_x j_x$,
4 of orders 1 in P and 3 in x,
$(12345)(12)3_x 4_x 5_x S$, $(12345)(23)1_x 5_x 4_x S$,
5 of orders 2 in P and 1 in x,
$(12345)(23)(45)1_x S$, $(12345)(43)(15)2_x S$, $(12345)(12)(45)3_x$,
5 of orders 2 in P and 3 in x,
$(12345)(21)(15)1_x 3_x 4_x S$, $(12345)(12)(23)2_x 4_x 5_x S$,
$(12345)(23)(34)3_x 1_x 5_x$,
20 of orders 3 in P and 1 in x,
$(12345)(12)(14)(15)3_x S$, $(12345)(21)(23)(25)4_x S$,
$(12345)(31)(32)(34)5_x$, $(12345)(12)(13)(45)1_x S$,
$(12345)(12)(15)(43)1_x S$, $(12345)(14)(15)(23)1_x S$,
$(12345)(21)(23)(54)2_x S$, $(12345)(24)(23)(51)2_x S$,
$(12345)(25)(21)(43)2_x S$, $(12345)(34)(35)(21)3_x S$,
$(12345)(34)(32)(15)3_x$.
2 of orders 3 in P and 3 in x, $(12345)(34)(32)(45)1_x 3_x 4_x S$.
18 of orders 4 in P and 3 in x,
$(12345)(ij)(ik)(im)(in)i_x$ 5 in number,
$(12345)(13)(15)(32)(34)1_x S$, $(12345)(14)(12)(43)(45)1_x S$,
$(12345)(15)(12)(53)(54)1_x$, $(12345)(21)(23)(14)(15)2_x S$,
$(12345)(24)(21)(43)(45)2_x S$, $(12345)(31)(34)(12)(15)3_x S$,
$(12345)(32)(34)(21)(25)3_x S$.
1 of orders 4 in P and 3 in x, $(12345)(12)(23)(34)(45)2_x 3_x 4_x$.
2 of orders 5 in P and 1 in x, $(12345)(31)(14)(23)(34)(45)1_x S$.

67 mixed forms containing p and u: These forms are the duals * of the mixed forms containing P and x and can be written down immediately. For example from the 5 forms

$$(12345)(ij)(ik)(im)(in)i_x,$$

we obtain the 5 dual forms

$$(kmn)(jmn)(jkn)(jkm)(jkmn).$$

* *Loc. cit.*, § 5.

In the above list i, j, k, m, n take the values 1, 2, 3, 4, 5 with the understanding that in any one form i, j, k, m, n are all distinct. The presence of the letter S after a form denotes the existence of a similar form,* that is a form in which the symbols 1 and 2 are interchanged with the symbols 5 and 4 respectively. To obtain the actual irreducible concomitants from this list we must remove from any form the invariant factors which appear. For example, $(12)^2 = a_\rho^2(AP)^2$ yields the actual concomitant $(AP)^2$.

3. *Determination of the Prepared System.* Since we are now considering two quadratics in n variables for the case $n = 5$, there are six invariants † and five quadratic covariants † i_x^2, $(i = 1, 2, 3, 4, 5)$. By theorem I every concomitant,‡ multiplied by a suitable invariant factor, can be expressed in terms of the symbolic factors,

$$i_x, \quad (ij), \quad (ijk), \quad (ijkm), \quad (12345) \qquad (i, j, k, m = 1, 2, 3, 4, 5).$$

We must now determine, if ever in forming these bracket factors, we have disturbed any of the invariant factors, which appear when 12, 23, 34, or 45 are convolved together. Originally we have five sets of cogredient point variables x, y, z, t, w, which are convolved as $\Delta = (xyztw)$, $u = xyzt$, $p = xyz$, $P = xy$. Since the only factor involving Δ is (12345) and since 12, 23, 34, 45, are all convolved in this, no invariant factor has been disturbed in forming it. When all the variables w have been convolved with $xyzt$ to form Δ, we are left to consider 4-factors, 3-factors and 2-factors, where an i-factor is a factor involving i of the variables x, y, z, t. We may neglect all 4-factors, as they lead to nothing new, for then the variables can only be $xyzt = u$. Let us now consider the possible cases, in which x, y, z, t may be convolved to form u. If one of these variables occur in a 3-factor, we may assume that three of them occur in this 3-factor, for $(ijk \mid x'zt)(rs \mid y'\xi) \equiv (ijkr' \mid xyzt)(s' \mid \xi) +$ terms in which xy are convolved together, and $(ijk \mid xy'z')(rs \mid t'\xi) \equiv (ijkr' \mid xyzt)(s' \mid \xi) +$ terms in which yzt are convolved together. Hence we must consider the cases when three variables occur in a 3-factor and the fourth occurs (a) in a 3-factor and (b) in a 2-factor.

Case (a) gives the possibility, $(ijkn' \mid u)(im' \mid \xi\eta)$, and case (b) $(ijkn' \mid u)(m' \mid \xi)$, where ξ, η may be any of x, y, z, t. In (a) and (b), neither of m, n is the same as any of i, j, k or else no convolution of successive integers is disturbed. At first sight it would appear that $(ijk'\dot{r}|u)(m'|\xi)(\dot{n}|\eta)$

* *Loc. cit.*, § 6.

† *Loc. cit.*, p. 404.

‡ *Loc. cit.*, § 3.

is a possibility, arising from three 2-factors, but i, j, k, r, m, n must all be distinct, and this is impossible. But if the variable t does not appear, we might have the single new type (c) $(ijk' \mid p)(m' \mid \xi)$, arising from two 2-factors.

We now write the factors for simplicity without the variables, since no confusion can arise. There are no further types of factors, as we shall see. Type (a) cannot occur with another u-factor, as

$$(ijkn')(im'rs), \quad (ijkn')(im'\dot{r}\dot{s})\dot{t}$$

are the only possibilities. In the first rs cannot contain i, m or n, and so must be jk. But by the fundamental identities this is impossible.* In the second case none of r, s, t can be i, therefore two of them must be either k, j or m, n and in either case no invariant factor is disturbed. Further since *

$$(ijkn')(irm') \equiv (ijkr)(imn) + (ijk)(inmr),$$

type (a) cannot occur with a further p-factor. Hence type (a) gives solely the one new factor type $(ijkn')(im')$.

Similarily it may be shown that type (b) cannot occur with another u or p-factor. Moreover type (c) cannot occur with another p-factor, for *

$$(ijk')(n'm) \equiv (ijm)(nk) + (ij)(knm),$$

and in both terms on the right i, j and n, k are convolved. There are three other possible cases to consider; $(ijkn')(m'a)$ from (b) and a 1-factor, $(ijkn')(m'\dot{a})(\dot{b}cde)$ from two (b) factors, and $(ijk')(m'ab)$ from one (c) factor and a 2-factor. Of these, the first reduces to type (a), since $a =$ one of i, j, k; the third gives nothing new, since one of a, b must be i or j; the second is more easily treated later.

In type (a), m, n must be consecutive integers and so must i, j. Hence we have the possibilities;

$$(1235')(14') \equiv (123, 154), \qquad (3215')(3'4') \equiv (123, 543),$$
$$(5431')(52') \equiv (543, 512).$$

In type (b) m, n must be consecutive integers and so we have;

$$1_x'(2'345), \quad 2_x'(3'145), \quad 3_x'(4'125), \quad 5_x'(4'321).$$

But

$$2_x'(3'145) \equiv 1_x(3245) + (3214')5_x' \equiv 5_x'(3214'),$$

since in (3245) both 2, 3 and 4, 5 are convolved. Similarly

$$3_x'(4'125) \equiv 1_x'(432'5).$$

* *Loc. cit.* Formulas (16) and (17).

Accordingly type (b) yields only two new factors,

$$1_x'(5432') \equiv (12, 543), \quad 5_x'(1234') \equiv (54, 123).$$

In type (c) both i, j and m, k must be successive integers and so we have,

$$1_x'(2'43) \equiv (12, 43), \quad 5_x'(4'23) \equiv (54, 23), \quad 1_x'(2'54) \equiv (12, 54).$$

If a new type of factor arises from two (b) factors, it must be

$$(1'\dot{4})(2'345)(\dot{5}321) \equiv (123'\dot{4})(4'5')(\dot{5}321),$$
$$\equiv (1254)(4'5')(3'321) \equiv (1254)(34')(5'321),$$

and so is expressible in terms of simpler factors. We thus have the eight new factors, three of type (a), three of type (c), and two of type (b). The factors of type (c) are the duals * of those of type (a), while each of the factors of type (b) is the dual of the other.

Now, since, with the addition of these new factor types, no invariant factors, which were originally introduced,† have been disturbed, we can work with the symbols i, j etc. and at the end remove all actual invariant factors and obtain the actual irreducible concomitants, provided that no identity is used, which separates successive integers convolved an even number of times. An alternative method is to use as a prepared system the factors (AP) for (12) etc. This prepared system was actually found by Dr. Wm. Saddler, but has never been published. He determined the prepared system by methods analogous to those used by H. W. Turnbull in his paper on two quadratics in four variables.‡ To find any of the irreducible concomitants by this method is cumbersome, as all identities have to be worked out in detail and in addition the ten symbols $a, r, A, R, A_3, R_3, \alpha, \rho$ must be paired off instead of the five symbols 1, 2, 3, 4, 5. This, together with the simplification of the identities, more than compensates for the addition of the extra factor (12345) and the fact that the identities cannot be applied blindly.

4. *Determination of the irreducible covariants and contravariants.* The factors which may occur in a covariant are the five i_x factors and (12345). The irreducible covariants are then six in number, the five quadratics i_x^2, and the quintic $(12345)1_x2_x3_x4_x5_x$. By duality § the contravariants are also six in number, the five quadratics $(kmnj)^2$ and the quintic

* *Loc. cit.*, § 5.

† *Loc. cit.*, p. 405.

‡ H. W. Turnbull, "The Simultaneous System of Two Quadratic Quaternary Forms," *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 18, Parts 1 and 2, pp. 70-94.

§ *Loc. cit.*, § 5.

$$(1234)(1235)(1245)(1345)(2345).$$

5. *Determination of the irreducible complexes.* The possible factors, which may occur, are the ten factors (ij) and (12345). But, as a product of (12345) by factors of the type (ij) always involves an odd number of symbols, the factor (12345) cannot appear in such a concomitant. Since the factors (ij) are strictly analogous to simple bracket factors of binary forms, we have only 20 possible complexes,

the 10 quadratics $(ij)^2$ and the 10 cubics $(ij)(jk)(ki)$,

for a product of four or more factors (ij) is reducible. In fact,

$$(ij)(km) \equiv (ik)(jm) + (kj)(im), \quad (ij)(kn) \equiv (ik)(jn) + (kj)(in).$$

By multiplying these two equations together and neglecting the terms, which involve a factor squared, we have

$$(ni)(ik)(kj)(jm) + (mi)(ik)(kj)(jn) \equiv 0,$$

or

$$2(ni)(ik)(kj)(jm) + (ji)(ik)(kj)(mn) \equiv 0,$$

by applying the identity $(mi')(j'n') \equiv 0$ to the second term. But as $(ji)(ik)(kj)$ is itself a concomitant $(ni)(ik)(kj)(jm)$ is reducible.*

By the principle of duality we see that there are only 20 irreducible complexes involving the variable p, the 10 quadratics $(ijk)^2$ and the 10 cubics $(12345)(ijk)(ijm)(ijn)$.

6. *Determination of the mixed concomitants containing u and x. Reductions.* (a) Since $(12, 543) \equiv 1_x'(5432') \equiv 3_x'(5'4'12)$, any concomitant containing the factor $(12, 543)$ is reducible, if 12 or both of 34, 45 are convolved an odd number of times. In addition such a concomitant has a factor $(AR_3)(AR_3ux)$ if 23 is convolved an odd number of times.

Further,

$$(12, 543) \equiv 3_x'(5'4'12) \equiv 3_x(5412) - 5_x'(34'12) \equiv 3_x(5412) - (54, 123).$$

Hence

(b) $\qquad (12, 543)(1254)3_x \equiv (12, 543)(54, 123) \equiv 0$, by (a).

There also exists a reduction similar to that for quaternary forms.†

(c) $\qquad (12, 543)(5432)2_x \equiv 0.$

---

* Grace and Young, *Algebra of Invariants*, Chap. 15, p. 322.

† J. Williamson, "Note on the Simultaneous System of Two Quadratic Quaternary Forms, *Journal of the London Mathematical Society*, Vol. 4 (1929), pp. 182-183.

For neglecting the invariant factors we have

$$(12, 543)(53432)2_x \equiv (AR_3ux)(A\rho x)u_\rho = 2(aR_3u)b_x[a_\rho b_x - a_x b_\rho]u_\rho,$$

and each term on the right has a factor b_x^2 or $b_\rho u_\rho b_x$. It is important to notice that the dual product $(54, 123)1_x(1345)$ is not reducible. Moreover

(d) $(1234)5_x M \equiv 0$, if 4, 5 is convolved an odd number of times in M, and $(2345)1_x M \equiv 0$, if 1, 2 is convolved an odd number of times in M.

Since $(12, 543) \equiv (5432)1_x - (5431)2_x$, by squaring this identity

(e) $$(5432)(5431)1_x 2_x \equiv 0.$$

If now we consider (5432) as simpler than (5431),

(f) $(5431)2_x M \equiv 0$, if 2, 3 is convolved an odd number of times in M.

Again by squaring the identity

$$(12, 543) - (2431)5_x \equiv (5231)4_x + (5421)3_x,$$

we have,

(g) $$(5231)(5421)4_x 3_x \equiv 0 \text{ by (a).}$$

In the above reductions we may replace each factor by its similar factor and in most cases obtain a new reduction. We now consider the possible forms in the following order; first those without the factor $(12, 543)$ and in ascending order in u, then those with one factor $(12, 543)$ and finally those with both factors $(12, 543)$ and $(54, 123)$. We only write down one of each pair of similar forms, and those forms which are marked R are reducible. The method of reduction is indicated shortly at the side.

One u factor. We have the five concomitants

$(ijkm)i_x j_x k_x m_x$,

and the types

$(12345)(1234)5_x$, $(12345)(1235)4_x\, R$(f), $(12345)(1245)3_x$.

Two u factors. We have the types

$(1234)(1235)4_x 5_x\, R$(e), $(1234)(1245)3_x 5_x\, R$(d),
$(1234)(2345)1_x 5_x\, R$(d), $(1234)(1345)2_x 5_x\, R$(d) and (f),
$(1235)(1245)3_x 4_x\, R$(g), $(1235)(1345)2_x 4_x\, R$(f) and (a),

and the ten $(12345)(ijkm)(ijkn)i_x j_x k_x$.

Three u factors. We have the ten $(mnjk)(mnij)(mnik)m_x n_x$, the duals of the previous case and the types

$(12345)(1245)(1345)(2345)1_x 2_x 3_x$,
$(12345)(1234)(1235)(1345)2_x 4_x 5_x$ R(f),
$(12345)(1423)(1425)(1435)2_x 3_x 5_x$ R(d),
$(12345)(1523)(1524)(1534)2_x 3_x 4_x$ R by $(1524)3_x$,
$(12345)(2314)(2315)(2345)1_x 4_x 5_x$ R(d),
$(12345)(2415)(2413)(2435)1_x 3_x 5_x$ R(d).

Four u factors. We have the types

$(1234)(2315)(1245)(1345)2_x 3_x 4_x 5_x$ R(e),
$(1234)(1235)(1245)(2345)1_x 3_x 4_x 5_x$ R(d),
$(1234)(1253)(1345)(2345)1_x 2_x 4_x 5_x$ R(f),

and the five $(12345)(ijkm)(ijkn)(ijmn)(ikmn)i_x$, the duals of the forms with one u and four x factors.

Five u factors. Either no x factors or five x factors occur. The first case has already been considered and in the second case all possible forms are reducible.

One factor $(12, 543)$. We have the simple forms

$(12345)(12, 543)$, $(12, 543)^2$ and two similar forms.

One further u factor. By (a) we see that there is only one possibility $(12, 543)(1245)3_x$ R(b).

Two further u factors. We have the types,

$(12, 543)(1234)(2345)2_x 3_x 4_x$ R(c),
$(12, 543)(1235)(2345)5_x 3_x 2_x$ R(c),
$(12, 543)(1245)(1345)1_x 4_x 5_x$,
$(12, 543)(12345)(1245)(1234)3_x 5_x$ R(d),
$(12, 543)(12345)(1245)(1235)3_x 4_x$ R(f) mod $(12, 543)(54, 123)$,
$(12, 543)(12345)(2345)(1345)1_x 2_x$ R(f) mod $(12, 543)^2$.

Three further u factors. We have the types

$(12, 543)(1234)(1235)(1245)3_x 4_x 5_x$ R mod $(12, 543)(54, 123)$,
$(12, 543)(1234)(2345)(1345)1_x 2_x 5_x$ R mod $(12, 543)^2$,
$(12, 543)(12345)(1234)(1235)(2345)2_x 3_x$ R(c),
$(12, 543)(12345)(1234)(1245)(1345)1_x 4_x$,
$(12, 543)(12345)(1235)(1245)(1345)1_x 5_x$.

Four further u factors. We have the sole possibility

$(12, 543)(12345)(1345)(2345)(1234)(1235)1_x 2_x 4_x 5_x$
R mod $(12, 543)^2$.

Five further u factors. There are no irreducible forms of this type. There are thus six irreducible forms containing one factor of the type $(12, 543)$, the

three in the list above and three similar forms. It is important to notice that each of these six is irreducible but that their duals reduce by (c).

Both factors (12, 543) *and* (54, 123). If both the factors (12, 543) and (54, 123) appear in a concomitant, since 12, 54, 23, and 34 must all be convolved an even number of times, there are very few possibilities and finally we are left with

$$(12, 543)(54, 123)(1235)(1245)(1345)1_x 5_x,$$

and its dual

$$(12, 543)(54, 123)(12345)(1234)(2345)2_x 3_x 4_x \quad R(c).$$

7. *Determination of the mixed concomitants containing P and x. Reductions.* If we consider the P factors in the order of simplicity,

(12), (54), (23), (34), (15), (14), (52), (13), (53), (24),

by identities of the type $(i'j')k_x' \equiv 0$, we see that

(h)

the products $(13)2_x$, $(53)2_x$, $(53)4_x$, $(13)4_x$, $(24)3_x$,
$(24)1_x$, $(25)1_x$, $(42)5_x$, $(41)5_x$,

are reducible. Further by identities of the type $(ij')(k'm') \equiv 0$, we see that the products

(i)

(13)(24), (53)(24), (13)(52), (53)(14), (14)(25)(15),
(24)(14)(15), (42)(52)(51), (14)(24)(25)

are reducible. The concomitant

(j) $$(13)(35)(51)M \equiv 0$$

also, since (13)(35)(51) is an actual concomitant containing no invariant factors.

We first consider those concomitants, which do not contain the factor (12345).

No factor (12345). In this case the only irreducible concomitants that appear are the ten mixed forms $(ij)i_x j_x$. This follows as a result of the analogy with binary forms (See § 5).

Forms containing the factor (12345). If the factor (12345) appears in a concomitant M, M must contain five x factors, three x factors or one x factor, since the number of symbols in a P factor is even. Five x factors cannot occur in M, for in that case the symbols appearing in the P factors must be paired

off. Accordingly M contains at least one factor (ij), in which i, j are not successive integers, and as a result the concomitant factor $(ij)i_x j_x$.

By considering list (i), we see that there can be no irreducible concomitants involving eight or more P factors. But if seven P factors occur and (24) occurs, (13) and (53) cannot appear nor can any of the products,

$$(14)(25)(15),\ (24)(15)(14),\ (24)(51)(52),\ (24)(14)(25)$$

and so in this case seven P factors cannot occur. But, if (24) does not occur, the products

$$(13)(52),\ (53)(14),\ (13)(35)(52),\ (14)(25)(15),\ (13)(35)(51)$$

are prohibited. Hence the only possible form involving seven P factors is $(12)(23)(34)(45)(25)(15)(35)M$ or the similar form. Since 1_x must occur among the x factors in M, this form *is* reducible by (h). Hence there are no irreducible concomitants containing seven P factors.

We shall now consider the remaining concomitants in ascending order in P. Since (12), (54); (13), (53); (23), (43); (25), (41); are similar factors and (15), (24) self similar factors, we need only write down one of every two similar forms.

One P factor. There is only one type, $(12345)(ij)k_x m_x n_x$ but all concomitants of this type are equivalent to

$$(12345)(12)3_x 4_x 5_x,\quad (12345)(23)1_x 4_x 5_x$$

and two similar forms by reductions (h).

Two P factors, one x factor. There is only one type $(12345)(ij)(km)n_x$. If we let $n = 1, 2, 3$ in succession and use reductions (h), we are left with

$$(12345)(23)(45)1_x,\ (12345)(43)(15)2_x,\ (12345)(12)(45)3_x,$$

and two similar forms.

Two P factors, three x factors. There is only one type $(12345)(ij)(ik)m_x n_x i_x$. Since, if i, k are not successive integers,

$$(12345)(ij)(ik)m_x n_x i_x \equiv (12345)(ij)(im)k_x n_x i_x \text{ etc.},$$

by letting $i = 1, 2, 3$ in turn we see that all forms of this type reduce to

$$(12345)(12)(15)1_x 3_x 4_x,\ (12345)(21)(23)2_x 4_x 5_x,\ (12345)(32)(34)3_x 1_x 5_x,$$

and two similar forms.

Three P factors, one x factor. There are two possible types

$$(1234)(ij)(ik)(im)n_x,\ (12345)(ij)(ik)(mn)i_x.$$

The concomitants of the first type reduce by reductions (h) to

$$(12345)(12)(14)(15)3_x,\ (12345)(21)(23)(25)4_x,\ (12345)(31)(32)(34)5_x,$$

and two similar forms. For, since the form

$$\begin{aligned} X &= (12345)(31)(32)(34)5_x, \\ &\equiv (12345)(35)(34)(32)1_x + (12345)(15)(32)(34)3_x = Y + Z, \end{aligned}$$

and the form Z appears in the list of irreducible forms of the second type, we may neglect Y, the form similar to X. The concomitants of the second type are equivalent to

$$\begin{array}{ll} (12345)(12)(13)(45)1_x, & (12345)(12)(15)(43)1_x, \\ (12345)(14)(15)(23)1_x, & (12345)(21)(23)(54)2_x, \\ (12345)(24)(23)(51)2_x, & (12345)(25)(23)(14)2_x \quad R, \\ (12345)(25)(21)(43)2_x, & (12345)(34)(35)(21)3_x, \\ (12345)(34)(32)(15)3_x, & \end{array}$$

and seven similar forms. The form marked R reduces by the identity $(25')(1'4') \equiv 0$.

Three P factors, three x factors. There is only one type

$$(12345)(mj)(ji)(ik)i_xj_xn_x.$$

In this i, j; i, k; j, m must all be successive integers, or else the form reduces by the identities $(i'j')n_x' \equiv 0$, $(i'k')j_x' \equiv 0$, $(j'm')i_x' \equiv 0$. Accordingly we are left with $(12345)(34)(32)(45)1_x3_x4_x$ and its similar form.

Four P factors, one x factor. There are two types $(12345)(ij)(ik)(im)(in)i_x$, $(12345)(ij)(ik)(jm)(jn)i_x$. The first type yields five concomitants. In the second type, if $i=1$ and $j=2$, one of (24) or (25) must occur and so the form is reducible. If $i=1$ and $j=3$, (35) cannot occur and so we have the possible form $(12345)(13)(15)(32)(34)1_x$. If $i=1$ and $j=4$, (42) cannot occur and if $i=1$ and $j=5$, (52) cannot occur and so we have the two forms $(12345)(14)(12)(43)(45)1_x$, $(12345)(15)(12)(43)(45)1_x$, the second of which is equivalent to its similar form. By a similar treatment for the cases $i=2$, and $i=3$, we have as a final list of concomitants of the second type

$$\begin{array}{ll} (12345)(13)(15)(32)(34)1_x, & (12345)(14)(12)(43)(45)1_x, \\ \multicolumn{2}{l}{(12345)(15)(12)(53)(45)1_x \equiv \text{to its similar form},} \\ (12345)(21)(23)(41)(15)2_x, & (12345)(24)(21)(43)(45)2_x, \\ (12345)(31)(34)(12)(15)3_x, & (12345)(32)(34)(21)(25)3_x, \end{array}$$

and six similar forms.

Four P factors, three x factors. There are three types

$$(12345)(mn)(ij)(jk)(ki)i_x j_x k_x, \quad (12345)(mi)(ij)(jk)(kn)i_x j_x k_x,$$
$$(12345)(mn)(nj)(jk)(kn)i_x j_x k_x.$$

The first type is not possible, since all of i, j; j, k; k, i cannot be successive integers and so the concomitant resolves into factors. In the second type m, i; i, j; j, k; k, n must be successive integers and so we have only the one irreducible form $(12345)(12)(23)(34)(45)2_x 3_x 4_x$. In the third type n, j; j, k; k, n must all be successive integers and this is impossible.

Five P factors, one x factor. There is only one type

$$(12345)(ij)(ik)(mj)(jk)(kn)i_x.$$

In such a concomitant, by the identity $(j'k')i_x' \equiv 0$, we see that j, k must be successive integers, and by identities of the type $(i'j')(k'n) \equiv 0$, that one factor of each of the products $(ij)(kn)$ and $(mj)(ik)$ must be a pair of successive integers. Since j, k are successive integers, both i, j and m, j cannot be successive integers. If i, j are successive integers, it follows from the above that i, k must be successive integers. But this is impossible since i, j; j, k; k, i cannot all be successive integers. Therefore i, j cannot be successive integers and so k, n must be. Since k, j are also successive integers, i, k cannot be successive integers and so m, j must be. As a result we have only two concomitants of this type $(12345)(13)(14)(23)(34)(45)1_x$ and the similar form.

Five P factors, three x factors. There are four types

$$(12345)(ij)(jk)(kn)(ni)(mn)i_x j_x k_x,$$
$$(12345)(in)(nk)(km)(mi)(mn)i_x j_x k_x,$$
$$(12345)(ij)(jk)(ki)(mi)(in)i_x j_x k_x,$$
$$(12345)(km)(mj)(jk)(mi)(in)i_x j_x k_x.$$

In the first type i, j; j, k; k, n; n, i must all be successive integers, and this is impossible. Similarly the last two types are also impossible. In the second type, if m, n are not successive integers, the form is reducible by the identity $(m'n')i_x' \equiv 0$. Therefore both of k, m and m, i cannot be successive integers. Further either i, n or k, m must be successive integers and also one of the pairs n, k and m, i. If i, n are successive integers, m, i cannot be, since m, n are successive integers, and therefore n, k must be successive integers while k, n cannot be. Hence the form has the factor $(km)(mi)k_x i_x$. If i, n are not successive integers, k, m must be and so m, i cannot be. But, since one of the pairs n, k and m, i must be successive integers, n, k must be suc-

cessive integers. Hence k, m; m, n; n, k must all be successive integers and this is impossible. Accordingly there are no irreducible concomitants containing five P factors and three x factors.

Six P factors and one x factor. There are two types

$$(12345)(jk)(jm)(jn)(km)(kn)(nm)i_x,$$
$$(12345)(ji)(ik)(mj)(jk)(km)(nm)i_x.$$

Of these two types we need only consider the second, for, in the first type one of k, m; k, n; n, m cannot be successive integers and so we can apply an identity of the type $(m'n')i_x' \equiv 0$ and reduce it to two forms of the second type. In the second type we see that k, j must be successive integers and that one factor of each of the products $(mj)(ik)$ and $(mk)(ij)$ must be a pair of successive integers. But it is impossible for this to be the case, since the product of (kj) with one of both pairs (mk), (ij) and (mj), (ik) consist of three factors with a symbol in common or else is of the type $(ij)(jk)(ki)$.

Six P factors and three x factors. There are three types

$$(12345)(im)(mj)(jk)(kn)(ni)(mn)i_xj_xk_x,$$
$$(12345)(ij)(jm)(mk)(ki)(mi)(in)i_xj_xk_x,$$
$$(12345)(im)(mj)(jn)(ni)(mk)(kn)i_xj_xk_x.$$

In the first type m, j; j, k; k, n must be successive integers. Accordingly m, n cannot be successive integers and the form reduces by $(m'n')i_x' \equiv 0$. In the second type m, i; m, j; m, k must all be successive integers and this is impossible. In the third type i, m; m, j; j, n; n, i cannot all be successive integers and so the form reduces to the first type by identities of the type $(i'm')k_x' \equiv 0$. Hence there are no irreducible concomitants containing six P factors. We have already shown that there are no irreducible concomitants with more than six P factors and so the list given in § 2 is complete.

By the principle of duality we can write down the irreducible concomitants involving the variables p and u.

The determination of the concomitants containing the variables P and x has not been attempted. The list of irreducible forms would be considerably longer but could be obtained by the methods that we have used. Since the complete system for two quadratics in four variables contains 122 forms, we should expect to obtain at least 700 or 800 forms in the complete system for two quadratics in five variables, as the labour involved in the latter case is at least five times as heavy as that in the former.

THE JOHNS HOPKINS UNIVERSITY.

Rational Surfaces Defined by Linear Systems of Plane Curves $C_{3n}: 8A^nB^{n-1}$.

By Joseph Crawford Polley.

1. *Introduction.* The rational surfaces of order four and five having no multiple curves, and those of order five having multiple curves insufficient for rationality, have been determined. There are three types of rational quartic surfaces with no double curve. One was discovered by L. Cremona * by applying a cremona transformation to a known quartic surface. The two remaining types were determined by M. Noether.† His method of investigation was that of considering quartic surfaces with a double point and applying to their equations the conditions for a one to one correspondence with the points of a plane.

Of particular interest also is the work of D. Montesano on rational quintic surfaces.‡ He obtained all the possible types by applying special cremona transformations to known rational surfaces.

In this paper various rational surfaces are discussed by considering certain linear systems of plane curves. The surfaces of Cremona, Noether and some of those of Montesano are re-determined by this method and a general type of surface is discussed by means of a linear system of plane curves of the form $C_{3n}: 8A^nB^{n-1}$.

2. *The system* $C_6: 7A^2$. In a plane (x) the system of curves $C_6: 7A^2$, with double points at 7 points A_i $(i = 1, 2, \cdots, 7)$, is of dimension 6. Let $C_3: 7A$, $C_3': 7A$ and $C_3'': 7A$ be linearly independent members of the net of cubics determined by the points A_i; and $C_6: 7A^2$ a non-composite sextic of the system. Then we can take as the equation of the system

$$(1) \quad a_1C_3^2 + a_2C_3C_3' + a_3C_3C_3'' + a_4C_3'C_3'' + a_5C_3'^2 + a_6C_3''^2 + a_7C_6 = 0.$$

Let

$$(2) \quad \begin{aligned} &y_1 = C_3^2: 7A, \quad y_2 = C_3: 7A \cdot C_3': 7A, \quad y_3 = C_3: 7A \cdot C_3'': 7A, \\ &y_4 = C_3': 7A \cdot C_3'': 7A, \quad y_5 = C_3'^2: 7A, \quad y_6 = C_3''^2: 7A, \quad y_7 = C_6: 7A^2. \end{aligned}$$

* L. Cremona, *Coll. Math. Chelini* 413-424 (1881).

† M. Noether, *Mathematische Annalen,* Vol. 33 (1889), pp. 546-571.

‡ D. Montesano, *Rendiconti della Reale Accademia di Napoli,* Ser. 3, Vol. 13 (1907), pp. 66-68.

These are the parametric equations of a rational surface F_a in S_6, in $(1,1)$ correspondence with the plane (x), the image of a point (x) being a point (y) on the surface. Since any two members of the set have eight residual intersections, a general S_4 in S_6 meets F_a in eight points. Hence F_a is of order eight.

For a general point on C_3

$$(3) \quad y_1 = y_2 = y_3 = 0, \quad y_4 = C_3'C_3'', \quad y_5 = C_3'^2, \quad y_6 = C_3''^2, \quad y_7 = C_6.$$

Hence the image of C_3 is a curve L in a sub-space S_3 of S_6.

Since a $C_6: 7A^2$ has, with $C_3: 7A$, four residual intersections, a plane

$$y_1 = y_2 = y_3 = \beta_4 y_4 + \beta_5 y_5 + \beta_6 y_6 + \beta_7 y_7 = 0$$

meets L in four points. Hence L is a quartic curve of genus 1.

By projection from the plane of $y_1 = y_2 = y_3 = y_7 = 0$ the surface F_8 goes into a surface F_4, in an S_3, whose parametric equations are

$$(4) \quad \begin{array}{ll} y_1 = C_3^2: 7A, & y_2 = C_3: 7A \cdot C_3': 7A, \\ y_3 = C_3: 7A \cdot C_3'': 7A, & y_7 = C_6: 7A^2. \end{array}$$

The image of L is the point $(0,0,0,1)$. To the plane sections of F_4 corresponds the system

$$a_1 C_3^2 + a_2 C_3 C_3' + a_3 C_3 C_3'' + a_4 C_6 = 0$$

all members of which pass through the four simple points in which C_3 meets C_6. The system is therefore of grade 4.

For a point near P_i $(i = 1, 2, 3, 4)$, the residual intersections of C_3 and C_6, the corresponding y_2, y_3 and y_7 are infinitesimals of the first order and the corresponding y_1 is an infinitesimal of the second order. Hence the images of P_i are straight lines in the plane $y_1 = 0$ through the point $(0,0,0,1)$.

The surface (4) is the well known rational quartic surface of Cremona.

3. *The system* $C_7: A^3 8B^2$. In a plane (x) the system of curves $C_7: A^3 8B^2$, with a triple point at A and double points at B_i $(i = 1, 2, \cdots, 8)$, is of dimension 5. The basis points determine a cubic $C_3: A8B$ and a web of quartics $C_4: A^2 8B$. Hence, if $C_7: A^3 8B^2$ and $C_7': A^3 8B^2$ are two non-composite curves of the system, and $C_1: A$, $C_1': A$ are members of the pencil of lines on A, the following can be taken as the linearly independent members of the system:

$$(5) \quad \begin{array}{lll} C_7: A^3 8B^2, & C_3: A8B \cdot C_4: A^2 8B, & C_3^2: A8B \cdot C_1: A, \\ C_3^2: A8B \cdot C_1': A, & C_3: A8B \cdot C_4': A^2 8B, & C_7': A^3 8B^2. \end{array}$$

A member of the pencil $C_7 + \gamma C_7' = 0$ has two residual intersections with C_3 and for a particular choice of γ there is a member tangent to C_3 at some point C. Let C_7' be that member and consider the pencil $C_7' + \delta C_3 C_4' = 0$. For a particular choice of δ we obtain a C_7 with a double point at C; call it $\bar{C}_7$. Furthermore, among the quartics there is a net through the point C, $C_4: A^2 8BC$, $C_3: A8BC \cdot C_1: A$ and $C_3: A8BC \cdot C_1': A$. Let D be the residual intersection of C_4 and C_3. Then we have as linearly independent members of the system

$$(6)\quad \begin{array}{lll} \bar{C}_7: A^3 8B^2C^2, & C_3: A8BC \cdot C_4: A^2 8BCD, & C_3^2: A8BCD \cdot C_1: AD, \\ C_3^2: A8BCD \cdot C_1': A, & C_3: A8BCD \cdot C_4': A^2 8B, & C_7': A^3 8B^2. \end{array}$$

Let

$$(7)\quad y_1 = \bar{C}_7,\ y_2 = C_3 \cdot C_4,\ y_3 = C_3^2 \cdot C_1,\ y_4 = C_3^2 \cdot C_1',\ y_5 = C_3 \cdot C_4',\ y_6 = C_7',$$

and project from the line $y_1 = y_2 = y_3 = y_4 = 0$ into the opposite S_3 of the S_5, thus obtaining a surface whose parametric equations are

$$(8)\quad \begin{array}{ll} y_1 = C_7: A^3 8B^2C^2, & y_2 = C_3: A8BCD \cdot C_4: A^2 8BCD, \\ y_3 = C_3^2: A8BCD \cdot C_1: 8AD, & y_4 = C_3^2: A8BCD \cdot C_1: 8A. \end{array}$$

The surface defined by equations (8) is a rational quartic surface F_4 since the system (8) is of grade 4.

For a general point on $C_3: A8BCD$, $y_1 \neq 0$ and $y_2 = y_3 = y_4 = 0$, hence the image of $C_3: A8BCD$ is the point $(1, 0, 0, 0)$ on F_4. The section of F_4 made by a plane $ky_2 + ly_3 = my_4 = 0$ through the point $(1, 0, 0, 0)$ determines in the plane (x) a composite curve

$$C_3: A8BCD(kC_4: A^2 8BCD + lC_3: A8BCD \cdot C_1: AD + mC_3: A8BCD \cdot C_1': A) = 0$$

that is

$$(9)\qquad C_3: A8BCD \cdot C_4: A^2 8BCD.$$

Any two curves of form (9) have two residual intersections. Therefore $(1, 0, 0, 0)$ is a double point. Since to each plane section through $(1, 0, 0, 0)$ corresponds a single C_4, hence but one direction through the point D in (x), any plane section through $(1, 0, 0, 0)$ has a cusp at that point.

Let a point P in (x) approach a point Q, other than D, on C_3; then y_3 and y_4 vanish to the second order, y_2 to the first order, and y_1 is finite; hence the image of P approaches $(1, 0, 0, 0)$ along $y_3 = y_4 = 0$, and $y_3 = y_4 = 0$ is the cuspidal tangent. As a point P approaches D on C_3, y_3 vanishes to the

third order, hence $(1, 0, 0, 0)$ is a uniplanar singular point on F_4, the plane of the point being the plane $y_3 = 0$.

The surface (8) is one of the rational quartic surfaces with no double curve, determined by Noether.

4. *The system* $C_9: 8A^3B^2$. In a plane (x) the system of curves $C_9: 8A^3B^2$ with triple points at A_i $(i = 1, 2, \cdots, 8)$ and a double point at B is of dimension 4. Taking $C_3: 8AB$ and $C_3: 8A$ as members of the pencil of cubics on A_i; $C_6: 8A^2B$, a non-composite sextic; and $C_9: 8A^3B^2$ a non-composite curve of order 9, the equation of the system is

$$\begin{aligned}(10) \qquad a_1C_9: 8A^3B^2 + a_2C_3: 8AB \cdot C_6: 8A^2B \\ + a_3C_3{}^2: 8AB \cdot C_3: 8A + a_4C_3{}^3: 8AB = 0.\end{aligned}$$

The cubic C_3 has with C_9 one residual intersection. Call this point C. Let

$$(11) \qquad \begin{aligned} y_1 &= C_9: 8A^3B^2C, & y_2 &= C_3: 8ABC \cdot C_6: 8A^2B, \\ y_3 &= C_3{}^2: 8ABC \cdot C_3: 8A, & y_4 &= C_3{}^3: 8ABC. \end{aligned}$$

These are the parametric equations of a rational surface of order 4 in ordinary space. The image of C_3 is the point $(1, 0, 0, 0)$.

For a point near C in (x) the corresponding values of y_1 and y_2 are infinitesimals of the first order and those of y_3 and y_4 are infinitesimals of the second order and the third order respectively. Hence the image of C is the line $y_3 = y_4 = 0$.

For a point near B in (x) the corresponding values of y_1, y_2 and y_3 are infinitesimals of the second order and that of y_4 is an infinitesimal of the third order. Hence the image of B is a conic in the plane $y_4 = 0$.

There is a pencil of curves $C_9: 8A^3B^2C^2$ given by

$$C_3{}^2: 8ABC \cdot C_3: 8A + \gamma C_3{}^3: 8ABC = 0.$$

These go into the sections of F_4 cut by the pencil of planes $y_4 = \gamma y_3$, these sections being composed of rational cubics and the line $y_3 = y_4 = 0$. To the section made by the plane $y_4 = 0$ corresponds $C_3{}^3: 8ABC$ which is of the form $C_9: 8A^3B^3C^3$, hence the section made by this plane is composed of a conic, image of B, and a line image of C, taken twice.

This surface is the second rational quartic surface with no double point determined by Noether.

5. *The system* $C_{12}: 8A^4B^3$. In a plane (x) the system of curves $C_{12}: 8A^4B^3$ is of dimension 4 and the equation of the system may be written

$$(12) \quad a_1C_{12}: 8A^4B^3D + a_2C_3: 8ABD \cdot C_9: 8A^3B^2 + a_3C_3^2: 8ABD \cdot C_6: 8A^2B + a_4C_3^3: 8ABD \cdot C_3: 8A + a_5C_3^4: 8ABD = 0$$

where D is the residual intersection of $C_3: 8AB$ and $C_{12}: 8A^4B^3$.

Let

$$y_1 = C_{12}: 8A^4B^3D, \quad y_2 = C_3: 8ABD \cdot C_9: 8A^3B^2,$$
$$y_3 = C_3^2: 8ABD \cdot C_6: 8A^2B, \quad y_4 = C_3^3: 8ABD \cdot C_3: 8A, \quad y_5 = C_3^4: 8ABD.$$

'hese are the parametric equations of a rational F_6 in S_4.

.et E be the residual intersection of $C_{12}: 8A^4B^3D = 0$ and $C_9: 8A^3B^2 = 0$. through E passes a curve of the pencil $C_3: 8A$ and a non-composite of the net $C_6: 8A^2B$. The image of E is a point P on F_6. Project ; from $(0, 0, 0, 0, 1)$ as a center into the opposite S_3 obtaining an F_5 parametric equations are

$$y_1 = C_{12}: 8A^4B^3DE, \qquad y_2 = C_3: 8ABD \cdot C_9: 8A^3B^2E,$$
$$y_3 = C_3^2: 8ABD \cdot C_6: 8A^2BE, \quad y_4 = C_3^3: 8ABD \cdot C_3: 8AE.$$

'ach C_{12} of the system goes into a plane section of F_5 which is a C_5 of 4, hence a C_5 with two double points. The locus of these double points be a double conic K on F_5.

The image of point D is the line $y_3 = y_4 = 0$; the image of point E is a line on the surface.

(15) The first polar of a rational surface F_N with respect to a point P not on the surface is a surface of order $N-1$, containing the double curve on F_N, the curve of contact of the tangent cone to F_N with P as a vertex, and the singular points on F_N. If F_N is a surface whose plane sections are mapped on a plane (x) by a web of curves of order m, then corresponding to the plane sections of F_N through P is a net of curves belonging to the web, whose Jacobian is a curve of order $3(m-1)$ having a $(3r-1)$-fold point at an r-fold point of the web. The Jacobian is the image of the curve of contact of the tangent cone with F_N.

For the case at hand the polar surface is an F_4 whose intersection with F_5 goes into a composite $C_{48}: 8A^{16}B^{12}D^4E^4$ consisting of the Jacobian $C_{33}: 8A^{11}B^8D^2E^2$; the curve $C_3: 8ABD$, image of $(0, 0, 0, 1)$; and the image of the double conic K which is a $C_{12}: 8A^4B^3DE^2$. A general $C_{12}: 8A^4B^3DE$ goes into a general plane section of F_5 but the $C_{12}: 8A^4B^3DE^2$ goes into the double conic counted twice and a residual line of F_5 in the plane of the conic. This line is the image of point E.

5(a). If point B is chosen on a certain locus there is a $C_9: 8A^3B^3$ other than

$C_3^3:8AB$.* Hence we can choose as linearly independent members of the system

$$(16)\quad \begin{array}{lll} C_9:8A^3B^3\cdot C_3:8ABC, & C_9:8A^3B^3\cdot C_3:8AC, & \\ C_3^3:8ABC\cdot C_3:8AC, & C_3^2:8ABC\cdot C_6:8A^2B, & C_3^4:8ABC, \end{array}$$

where C is the ninth point common to $C_3:8ABC$ and $C_3:8A$.

Take any point D in the plane (x). There is a member of th
$C_9:8A^3B^3+\gamma C_3^3:8ABC=0$ through D. Call this a new $C_9:8A^3B^3$.
is also a $C_3:8A$ and a $C_6:8A^2B$ through D.

Let

$$(17)\quad \begin{array}{ll} y_1=C_9:8A^3B^3D\cdot C_3:8ABC, & y_2=C_9:8A^3B^3D\cdot C_3:8ACD, \\ y_3=C_3^2:8ABC\cdot C_6:8A^2BD, & y_4=C_3^3:8ABC\cdot C_3:8ACD. \end{array}$$

These are the parametric equations of a rational surface F_5 in S_3.†

The image of $C_9:8A^3B^3D$ is the line $y_1=y_2=0$. The in
$C_3:8ACD$ is the line $y_2=y_4=0$. The image of the point C is
$y_3=y_4=0$, the image of point D a line in the plane $y_2=0$, and th
of point B a rational cubic in the plane $y_1=0$, with a double p
$(0,1,0,0)$.

Since C_9 has two residual intersections with a general C_{12}, eac
section of F_5 has a double point on the line $y_1=y_2=0$, the image
Hence $y_1=y_2=0$ is a double line.

A section of F_5 made by a plane of the pencil $y_2=\gamma y_4$ goes into a composite C_{12} of the form $C_3:8ACD\cdot C_9:8A^3B^3$, which meets a general C_{12} in 3 residual points not on $C_3:8ACD$. Hence the line of intersection of any plane of the pencil and a plane not of the pencil meets F_5 in 3 points other than on the line $y_2=y_4=0$, which means that $y_2=y_4=0$ is a double line.

We observe that the surface F_5 as defined by (17) has a composite double conic consisting of the double lines $y_1=y_2=0$ and $y_2=y_4=0$. As in the general case (14), the image line of the point D is the residual intersection of the plane of the double conic with F_5.

6. *The system* $C_{15}:8A^5B^4$. In a plane (x) the system of curves $C_{15}:8A^5B^4$ is of dimension 5 and by the method employed in the previous cases we can choose as linearly independent members of the system

$$(18)\quad \begin{array}{lll} C_{15}:8A^5B^4, & C_{12}:8A^4B^3\cdot C_3:8AB, & C_9:8A^3B^2\cdot C_3^2:8AB, \\ C_6:8A^2B\cdot C_3^3:8AB, & C_3^4:8AB\cdot C_3:8A, & C_3^5:8AB. \end{array}$$

* Halphen, *Bulletin de la So. Math.*, 162 (1882).

† D. Montesano, *Rendiconti della Reale Accodemia di Napoli*, Ser. 3, Vol. 13 (1907), pp. 66-68.

Let C be the residual intersection of $C_3: 8AB$ and $C_{15}: 8A^5B^4$. The curves $C_{15}: 8A^5B^4$ and $C_{12}: 8A^4B^3$ have 8 residual points of intersection. Call two of these points D and E. Choose a new C_9 and a new C_6 such that they will pass through D and E.

Let

$$(19)\quad \begin{array}{ll} y_1 = C_{15}: 8A^5B^4CDE, & y_2 = C_{12}: 8A^4B^3DE \cdot C_3: 8ABC, \\ y_3 = C_9: 8A^3B^2DE \cdot C_3^2: 8ABC, & y_4 = C_6: 8A^2BDE \cdot C_3^3: 8ABC, \\ y_5 = C_3^4: 8ABC \cdot C_3 8AD, & y_6 = C_3^5: 8ABC. \end{array}$$

These are the parametric equations of a rational F_8 in S_5. The images of points D and E are points on the line $y_1 = y_2 = y_3 = y_4 = 0$.

From the line $y_1 = y_2 = y_3 = y_4 = 0$ as a center project F_8 into the S_3, $y_5 = y_6 = 0$, giving a surface whose parametric equations are

$$(20)\quad \begin{array}{ll} y_1 = C_{15}: 8A^5B^4CDE, & y_2 = C_{12}: 8A^4B^3DE \cdot C_3: 8ABC, \\ y_3 = C_9: 8A^3B^2DE \cdot C_3^2: 8ABC, & y_4 = C_6: 8A^2BDE \cdot C_3^3: 8ABC. \end{array}$$

The surface is an F_6. The points D and E go into lines on this surface. The image of C is the line $y_3 = y_4 = 0$. The image of $C_3: 8ABC$ is the point $(1, 0, 0, 0)$.

Since the genus of a member of the system is 5, each plane section of F_6 is a C_6 of genus 5. Hence there is a double C_5 on the surface.

By (15) we see that the section of F_6 by the first polar of a point not on F_6 goes into a composite $C_{75}: 8A^{25}B^{20}C^5D^5E^5$ which consists of the Jacobian $C_{42}: 8A^{14}B^{11}C^2D^2E^2$; $C_3: 8ABC$, the image of the singular point $(0, 0, 0, 1)$; and the image of the double C_5 which is a $C_{30}: 8A^{10}B^8C^2D^3E^3$. A general $C_{30}: 8A^{10}B^8C^2D^2E^2$ goes into a general quadric section of F_6 but the $C_{30}: 8A^{10}B^8C^2D^3E^3$ goes into the quadric section containing the double C_5 with the image lines of points D and E as the residual intersection of the quadric with F_5.

If two surfaces of order n_1 and n_2 respectively contain a C_m of order m and rank r, genus p to multiplicity i_1 and i_2 respectively, the residual C_v meets C_m in t points and has genus π where *

$$(21)\quad \begin{aligned} t &= m(i_2n_1 + i_1n_2 - 2i_1i_2) - i_1i_2r, \\ \pi &= [v(n_1 + n_2 - 4) - (i_1 + i_2 - 1)t]/2 + 1, \\ r &= 2m + 2p - 2. \end{aligned}$$

* Noether, *Annali di Mathematica*, Ser. 2, Vol. 5 (1871), pp. 163-177.

For the case in question

$$\begin{array}{lll} n_1 = 2 & i_1 = 1 & m = 5 \\ n_2 = 6 & i_2 = 2 & \end{array}$$

and, since C_v is a composite conic,

$$v = 2 \text{ and } \pi = -1.$$

Substituting the above values in equations (21) we find that $p = 2$; that is, the double C_5 on F_6 is a curve of genus 2.

A C_5 genus 2 is the partial intersection of a quadric and a cubic surface, the residual being a ruling on the quadric.

Let F_3 be a cubic surface containing the C_5 and one line of the degenerate C_2. Then the number of intersections of the line with C_5 is, by (21), $t = 3$. Hence the line images of points D and E are trisecants of the quintic C_5.

6(a). If B is chosen on a certain locus there is a $C_{12}: 8A^4B^4$ other than $C_3{}^4: 8AB$, and we can take as parametric equations of the surface

$$(22)\quad \begin{array}{ll} y_1 = C_{12}: 8A^4B^4DE \cdot C_3: 8AC, & y_2 = C_{12}: 8A^4B^4DE \cdot C_3: 8ABC, \\ y_3 = C_9: 8A^3B^2DE \cdot C_3{}^2: 8ABC, & y_4 = C_6: 8A^2BDE \cdot C_3{}^3: 8ABC. \end{array}$$

The surface is again an F_6 with plane sections of genus 5 and a double curve C_5 of order 5.

The image of C_{12} is the line $y_1 = y_2 = 0$. Since $C_{12}: 8A^4B^4DE$ meets a general $C_{15}: 8A^5B^4CDE$ in 2 residual points, $y_1 = y_2 = 0$ is a double line; hence the double C_5 is composite with $y_1 = y_2 = 0$ as a component.

For a point near B in (x), y_1, y_3 and y_4 are infinitesimals of order four while y_2 is an infinitesimal of order five, hence the image of B is a rational quartic in the plane $y_2 = 0$. The images of D and E are lines meeting the line $y_1 = y_2 = 0$, since D and E are on C_{12} in plane (x).

6(b). Again if B is chosen on a certain locus so that there is a $C_9: 8A^3B^3$ other than $C_3{}^3: 8AB$ we can take the following as parametric equations of a surface

$$(23)^*\quad \begin{array}{l} y_1 = C_9: 8A^3B^3DEF \cdot C_6: 8A^2BCDEF, \\ y_2 = C_9: 8A^3B^3DEF \cdot C_3: 8ABC \cdot C_3: 8A, \\ y_3 = C_9: 8A^3B^3DEF \cdot C_3{}^2: 8ABC, \\ y_4 = C_6: 8A^2BDEF \cdot C_3{}^3: 8ABC. \end{array}$$

The surface is an F_5 genus 5. Since each plane section is a quintic curve of genus 5 there is a double line on the surface.

* D. Montesano, *Rendiconti della Reale Accademia di Napoli*, Ser. 3, Vol. 13 (1907), pp. 66-68.

The image of C_9 is the point $(0,0,0,1)$. The image of $C_3: 8AB$ is the point $(1,0,0,0)$. The image of C is the line $y_3 = y_4 = 0$. The images of points D, E and F which are on C_9 are three lines in the plane $y_1 = 0$, passing thru $(0,0,0,1)$, the image of C_9.

The image of C_6 is the line $y_1 = y_4 = 0$. Since $C_6: 8A^2BCDEF$ has, with a general $C_{15}: 8A^5B^4CDEF$, two residual intersections, every plane section of F_5 has a double point on $y_1 = y_4 = 0$, which is, then, the double line on the surface.

A plane section thru the point $(0,0,0,1)$ goes into a composite curve in plane (x) of the form

$$C_9: 8A^3B^3DEF \cdot C_6: 8A^2BC.$$

Two curves of this type meet in two residual points not on C_9, hence the line of intersection of two planes through $(0,0,0,1)$ meets F_5 in two points other than $(0,0,0,1)$, hence $(0,0,0,1)$ is a triple point on F_5.

By the method of (15) we can show that the images of the points D, E and F in (x) form the residual intersection of the plane through the double line and the point $(0,0,0,1)$.

7. *The system* $C_{18}: 8A^6B^5$. This system is of dimension 6 and by a process of reasoning similar to that in the previous cases we obtain the surface in S_3 given parametrically by

$$(24)\quad \begin{array}{ll} y_1 = C_{18}: 8A^6B^5CD^2, & y_2 = C_{15}: 8A^5B^4D^2 \cdot C_3: 8ABC, \\ y_3 = C_{12}: 8A^4B^3D^2 \cdot C_3^2: 8ABC, & y_4 = C_9: 8A^3B^2D^2 \cdot C_3^3: 8ABC, \end{array}$$

The surface is an F_6 and each plane section is a C_6 genus 5; hence the surface contains a double C_5 of order 5. The image of the curve C_3 is the point $(1,0,0,0)$. The image of the point C is the line $y_3 = y_4 = 0$. The image of the point D is a conic on F_6.

Through the double C_5 on F_6 passes one quadric surface whose residual intersection with F_6 is a conic C_2.

Again referring to (15) we can show that the double curve C_5 goes into a $C_{36}: 8A^{12}B^{10}C^2D^5$. Since a general $C_{36}: 8A^{12}B^{10}C^2D^4$ goes into a general quadric section of F_6, the $C_{36}: 8A^{12}B^{10}C^2D^5$ goes into the section made by the quadric on the double C_5, which, therefore, has the image conic of point D as the residual intersection with F_6.

7*(a)*. If B is chosen so that there is a $C_9: 8A^3B^3$ other than $C_3^3: 8AB$ we obtain a surface whose parametric equations are

$$(25)^* \quad \begin{aligned} y_1 &= C_9: 8A^3B^3DE \cdot C_9: 8A^3B^2CD, \\ y_2 &= C_9: 8A^3B^3DE \cdot C_6: 8A^2BD \cdot C_3: 8ABC, \\ y_3 &= C_9: 8A^3B^3DE \cdot C_3: 8AD \cdot C_3^2: 8ABC, \\ y_4 &= C_{12}: 8A^4B^3D^2E \cdot C_3^2: 8ABC. \end{aligned}$$

The surface is an F_5 with plane sections of genus 5; hence there is a double line on the surface. The image of $C_9: 8A^3B^3$ is the point $(0, 0, 0, 1)$. The image of $C_3: 8ABC$ is the point $(1, 0, 0, 0)$. The images of D and E are a conic and a line, both of which must pass through the point $(0, 0, 0, 1)$, since D and E are on $C_9: 8A^3B^3$. The image of C is the line $y_3 = y_4 = 0$.

The plane determined by the double line and the point $(0, 0, 0, 1)$ has a C_3 as residual intersection with F_5.

From (15) we find that the double line goes into a $C_{18}: 8A^6B^5CD^3E^2$. Since a general C_{18} goes into a general plane section of F_5, the $C_{18}: 8A^6B^5CD^3E^2$ goes into the section made by the plane through the double line and the point $(0, 0, 0, 1)$ and contains the images of points D and E as the residual intersection with F_5. The residual C_3 in which the plane determined by the double line and the point $(0, 0, 0, 1)$ meets the F_5 is therefore composite, and consists of a conic and a line, images of D and E.

A section of F_5 by a plane $a_1y_1 + a_2y_2 + a_3y_3 = 0$ goes into a composite curve of the form

$$C_9: 8A^3B^3DE(a_1C_9: 8A^3B^2CD + a_2C_6: 8A^2BD \cdot C_3: 8ABC \\ + a_3C_3: 8AD \cdot C_3: 8ABC) = 0;$$

that is

$$C_9: 8A^3B^3DE \cdot C_9: 8A^3B^2CD.$$

Two of these $C_9: 8A^3B^2CD$ meet in three residual points, hence $(0, 0, 0, 1)$ is a double point on F_5.

7(b). If the point B is so chosen that there is a $C_{15}: 8A^5B^5$ other than $C_3^5: 8AB$, the parametric equations of the surface may be taken as

$$(26) \quad \begin{aligned} y_1 &= C_{15}: 8A^5B^5DEF \cdot C_3: 8AC, & y_2 &= C_{15}: 8A^5B^5DEF \cdot C_3: 8ABC, \\ y_3 &= C_{12}: 8A^4B^3DEF \cdot C_3^2: 8ABC, & y_4 &= C_9: 8A^3B^2DEF \cdot C_3^3: 8ABC. \end{aligned}$$

This surface differs from (24) only in that the line $y_1 = y_2 = 0$ is a double line on the surface and a component of the double C_9.

7(c). If the point B is so chosen that there is a $C_{12}: 8A^4B^4$ other than $C_3^4: 8AB$ then we obtain a surface whose parametric equations are

* D. Montesano, *Rendiconti della Reale Accademia di Napoli*, Ser. 3, Vol. 13 (1907), pp. 66-68.

$$(27)\quad \begin{aligned} y_1 &= C_{12}: 8A^4B^4DEFG \cdot C_6 8A^2BC, \\ y_2 &= C_{12}: 8A^4B^4DEFG \cdot C_3: 8ABC \cdot C_3: 8A, \\ y_3 &= C_{12}: 8A^4B^4DEFG \cdot C_3^2: 8ABC, \\ y_4 &= C_9: 8A^3B^2DEFG \cdot C_3^3: 8ABC. \end{aligned}$$

The surface is an F_6 with plane sections of genus 6 and contains a double C_4. The image of C_{12} is the point $(0, 0, 0, 1)$. A section made by a plane through this point goes into a curve in plane (x) of the form

$$C_{12}: 8A^4B^4DEFG \cdot C_6: 8A^2BC.$$

Two of these $C_6: 8A^2BC$ meet in two residual points, hence the point $(0, 0, 0, 1)$ is 4-fold on the surface.

The images of D, E, F, and G are lines in the plane $y_1 = 0$ passing through $(0, 0, 0, 1)$. The image of C_6 is the line $y_1 = y_4 = 0$. In plane (x), $C_6: 8A^2BCDEFG$ meets a general $C_{18}: 8A^6B^5CDEFG$ in two residual points, hence $y_1 = y_4 = 0$ is a double line and a component of the double C_4.

From (15) we find that the double C_4 goes into a $C_{36}: 8A^{12}B^{10}C^2D^3E^3F^3G^3$. This C_{36} goes into the quadric section on the double C_4 and the point $(0, 0, 0, 1)$ which has as residual intersection with F_6 the composite C_4 consisting of the four lines, images of the points D, E, F and G.

8. *The system* $C_{21}: 8A^7B^6$. This system is of dimension 7 and by the methods previously employed we obtain the surface whose parametric equations are

$$(28)\quad \begin{aligned} y_1 &= C_{21}: 8A^7B^6CD^2E, & y_2 &= C_{18}: 8A^6B^5D^2E \cdot C_3: 8ABC, \\ y_3 &= C_{15}: 8A^4B^3D^2E \cdot C_3^2: 8ABC, & y_4 &= C_{12}: 8A^3B^2D^2E \cdot C_3^3: 8ABC. \end{aligned}$$

The surface is an F_7 with plane sections of genus 6, hence there is a double curve C_9 of order 9 on the surface.

The image of C is the line $y_3 = y_4 = 0$, the image of D a conic and the image of E a line. By (15) we find that the double C_9 goes into a curve in the plane (x) of the form $C_{63}: 8A^{21}B^{18}C^3D^7E^3$. This C_{63} goes into the section of F_7 made by the cubic surface containing the double C_9 and has as residual intersection a composite cubic curve consisting of the conic image of D and the line image of E.

8(a). If B is so chosen that there is a $C_{15}: 8A^5B^5$ other than $C_3^5: 8AB$ a surface is determined whose parametric equations are

$$(29)\quad \begin{aligned} y_1 &= C_{15}: 8A^5B^55D \cdot C_6: 8A^2BC, & y_2 &= C_{15}: 8A^5B^55D \cdot C_3: 8ABC \cdot C_3 8A, \\ y_3 &= C_{15}: 8A^5B^55D \cdot C_3^2: 8ABC, & y_4 &= C_{12}: 8A^4B^35D \cdot C_3^3: 8ABC. \end{aligned}$$

The surface is an F_7 with plane sections of genus 7 hence there is a double curve C_8 of order 8 on the surface. The image of C is the line $y_3 = y_4 = 0$. The image of $C_3 : 8ABC$ is the point $(0, 0, 0, 1)$. Since a line through $(0, 0, 0, 1)$ meets the surface in only two residual points this point is of order 5 on the surface. The image of B is a rational sextic in the plane $y_3 = 0$ with a point of order 5 at $(0, 0, 0, 1)$.

The images of the D_i $(i = 1, 2, \cdots, 5)$ are five lines on the surface passing through the point $(0, 0, 0, 1)$ and form the residual intersection with F_7 of the cubic surface through the double C_8.

8(b). If the point B is so chosen that there is a $C_{12} 8A^4B^4$ other than $C_3{}^4 : 8AB$ a surface is determined whose parametric equations are

$$(30) \qquad \begin{aligned} y_1 &= C_{12} : 8A^4B^4DEF \cdot C_9 : 8A^3B^2CD^2EF, \\ y_2 &= C_{12} : 8A^4B^4DEF \cdot C_6 : 8A^2BD \cdot C_3 : 8ABC, \\ y_3 &= C_{12} : 8A^4B^4DEF \cdot C_3{}^2 : 8ABC \cdot C_3 : 8AD, \\ y_4 &= C_9 : 8A^3B^2CD^2EF \cdot C_3{}^4 : 8ABC. \end{aligned}$$

The surface is an F_6 with plane sections of genus 6, hence there is a double C_4 on the surface.

The image of C_{12} is the point $(0, 0, 0, 1)$, the image of $C_3 : 8ABC$ the point $(1, 0, 0, 0)$, and the image of C_9 is the line $y_1 = y_4 = 0$. The image of D is a conic in the plane $y_1 = 0$ passing through the point $(0, 0, 0, 1)$. The images of E and F are lines in the plane $y_1 = 0$. The line $y_1 = y_4 = 0$ is a double line. The point $(0, 0, 0, 1)$ is a triple point on the surface.

By the methods employed in the previous cases we find that the residual intersection with F_6 of the quadric through the double C_4 is a composite C_4 consisting of the conic image of the point D and the line images of the points E and F.

9. *Conclusion.* It is now clear that the processes developed in this paper can be carried on indefinitely for any linear system of the type $C_{3n} : 8A^nB^{n-1}$ containing a pencil $C_{3i} : 8A^iB^i$.

A Problem of Ambience.

By William Kelso Morrill.

In the following paper, we shall consider a triangle of directed lines, the vertices of which are moving with the same constant speed parallel to their respective opposite sides but in opposite directions. The invariants of the triangle are studied, and the motions of the vertices are investigated by the aid of the Weierstrass elliptic function theory as well as the q-series of Jacobi.

1. *The Invariants of the Triangle.* Let a, b, c be the lengths of the sides of the triangle, and α, β, γ their respective directions. If θ, ϕ, ψ are the angles which the sides of the triangle make with the base line, then $-\alpha = e^{i\theta}$; $\beta = e^{i\phi}$; $-\gamma = e^{i\psi}$. We shall use A, B, C in two senses: first as the affices of the vertices, second as the interior angles of the triangle. The motion of the vertices is given by the following diffrential equations

$$\dot{A} = -\alpha v, \quad \dot{B} = -\beta v, \quad \dot{C} = -\gamma v,$$

where the dot indicates differentiation with respect to the time and v is the speed. We can now write $a\alpha = C - B$. Then

$$D_t a\alpha \equiv \dot{a}\alpha + \dot{\alpha}a = \dot{C} - \dot{B} = (\beta - \gamma)v$$

and
$$\dot{a} + (\dot{\alpha}/\alpha)a = (\beta/\alpha - \gamma/\alpha)\, v.$$

But
$$\beta/\alpha = e^{i(\pi - C)} = -\cos C + i \sin C$$

and
$$\gamma/\alpha = e^{i(\pi + B)} = -\cos B - i \sin B.$$

$$\therefore \quad \dot{a} + a\dot{\alpha}/\alpha = v[\cos B - \cos C + i(\sin B + \sin C)].$$

Equating reals and imaginaries,

$$\dot{a} = (\cos B - \cos C)v,$$
$$a\dot{\alpha}/\alpha = i(\sin B + \sin C)v.$$

The variations of the sides of the triangle are given, therefore, by

$$1.1 \quad \begin{cases} \dot{a} = (\cos B - \cos C)v, \\ \dot{b} = (\cos C - \cos A)v, \\ \dot{c} = (\cos A - \cos B)v. \end{cases}$$

Now $v = d\Lambda/dt$, where Λ is the distance each body moves in the time t. Choosing $v = 1$, we have $d\Lambda = dt$ and $\Lambda = t$. Adding equations 1.1, we have $\dot{a} + \dot{b} + \dot{c} = 0$.

1.2 $$\therefore \quad a + b + c = s_1$$

where s_1 is a constant; and our first result is that the perimeter of the triangle remains constant. Finding the perimeter constant suggests a study of the area.

$$\text{Area} = [s(s-a)(s-b)(s-c)]^{1/2},$$

where $s = (a+b+c)/2$. Thus, letting

$$X \equiv 16(\text{Area})^2 = 2(a^2b^2 + b^2c^2 + a^2c^2) - a^4 - b^4 - c^4$$
$$\dot{X} = 4\{(b^2 + c^2 - a^2)\dot{a}a + (c^2 + a^2 - b^2)\dot{b}b + (a^2 + b^2 - c^2)\dot{c}c\}$$
$$= 8abc(\dot{a}\cos A + \dot{b}\cos B + \dot{c}\cos C).$$

Substituting the values of $\dot{a}$, $\dot{b}$, $\dot{c}$ from 1.1; we have: $\dot{X} = 0$ and

1.3 $X = S_3$, where S_3 is a constant; that is, the area also is constant.

We thus find *the perimeter and area of the triangle are invariant under the motion.*

2. *Introducing the Elliptic Functions.* Let $x = s - a$, $y = s - b$, $z = s - c$. Then from 1.2 and 1.3 respectively, we obtain

$$x + y + z = k_1, \quad \text{and}$$
$$xyz = k_3.$$
$$\therefore \quad dx + dy + dz = 0, \quad \text{and}$$
$$yzdx + xzdy + xydz = 0.$$

From these two equations, we obtain

2.1 $$dx/x(y-z) = dy/y(z-x) = dz/z(x-y) = d\mu,$$
$$\therefore dz/d\mu = z(x-y) = z\{(x+y)^2 - 4xy\}^{1/2} = z[(k_1 - z)^2 - 4k_3/z]^{1/2}$$
$$(dz/d\mu)^2 = z^2[(k_1 - z)^2 - 4k_3/z].$$

Put $z = -1/v$; then $dz/d\mu = v^{-2}dv/d\mu$, and

$$(dv/d\mu)^2 = (k_1 v + 1)^2 + 4k_3 v^3.$$

Now putting $v = w - k_1^2/12k_3$, we get

2.2 $$(dw/d\mu)^2 = 4k_3 w^3 + (2k_1 - k_1^4/12k_3)w + (1 - 2k_1^3/12k_3 + 2k_1^6/3 \cdot 12^2 k_3^2).$$

Finally putting $k_3^{1/2} d\mu = du$, and

2.3 $$g_2 = (-2k_1/k_3)(1 - k_1^3/24k_3) \quad \text{and}$$

2.4 $$g_3 = (-1/k_3)(k_1^6/216k_3^2 - k_1^3/6k_3 + 1) \quad \text{and}$$

substituting in 2. 2 we obtain

2. 5 $$(dw/du)^2 = 4w^3 - g_2w - g_3.$$

Equation 2. 5 is the elliptic relation $(p'u)^2 = 4p^3u - g_2pu - g_3$, and hence our problem is an elliptic function problem.

We may then write $w = p(u - \gamma)$, where γ is a constant. Since $z = -1/v$, and $v = w - k_1^2/12k_3$, $z = 1/[k_1^2/12k_3 - p(u - \gamma)]$. Setting

2. 6 $$k_1^2/12k_3 = p\tau$$

and noting, from 2. 1, that x and y have expressions similar to z, we have:

2. 7 $$\begin{cases} x = 1/[p\tau - p(u - \alpha)], \\ y = 1/[p\tau - p(u - \beta)], \\ z = 1/[p\tau - p(u - \gamma)]. \end{cases}$$

Note that $z^{-1}dz/du = (x - y)/k_3^{1/2} = p'(u - \gamma)/[p\tau - p(u - \gamma)]^2$. This equals zero when $x = y$, which is at the half periods of the parallelogram of periods, since we know the function p' is zero there. Conversely when $u - \gamma$ is a half period, that is when

2. 8 $$u - \gamma \equiv m_1\omega_1 + m_2\omega_2,$$ where

m_1, $m_2 = 0$, 1, 2 but $m_1 = m_2 \neq 0$ or 2, $p'(u - \gamma) = 0$, and $x = y$ or from 2. 7, $p(u - \alpha) = p(u - \beta)$.

Consider then $u - \alpha \equiv \beta - u$, and, therefore, $2u \equiv \alpha + \beta$. Since from 2. 8 we have $2u \equiv 2\gamma$, it follows that

2. 9 $$\alpha + \beta \equiv 2\gamma, \quad \beta + \gamma \equiv 2\alpha, \quad \gamma + \alpha \equiv 2\beta.$$
$$\therefore \quad \alpha - \gamma \equiv 2\gamma - 2\alpha \quad \text{or} \quad 3\alpha \equiv 3\beta \equiv 3\gamma.$$

This result tells us that α, β, γ in 2. 7 are constants which differ from each other by thirds of a period. By choosing $\gamma = 0$, it follows that $\alpha = \alpha$, and $\beta = 2\alpha$, and we can rewrite 2. 7 in the following way:

2. 7′ $$\begin{cases} x = 1/(p\tau - pu), \\ y = 1/[p\tau - p(u + \alpha)], \\ z = 1/[p\tau - p(u - \alpha)]. \end{cases}$$

x has poles at $u = \pm\tau$; y has poles at $u = \pm\tau - \alpha$; and z has poles at $u = \pm\tau + \alpha$. Since $x + y + z$ is a constant, the sum of the poles of x, y, and z respectively must be a period. Hence τ must be a third of a period.

The type of network can now be determined quite easily: g_2 and g_3 are both real and the discriminant

$$\Delta \equiv g_2^3 - 27g_3^2 = (k_1^3 - 27k_3)/k_3^3 > 0.$$

This follows from the theorem: *If n numbers $x_1, \cdots x_n$ are positive, the arithmetical mean must be equal to or greater than the geometrical mean.** Since g_2 and g_3 are real and $\Delta > 0$, our net work is rectangular.

We are interested in how the triangle behaves as the elliptic parameter u moves in a rectangular cell. But there are limitations on how u shall move in the cell.

There are eight thirds of a period in a cell. Of these, only four give distinct values to pu, due to the evenness of the p function. At the vertices of the cell pu is infinite. Along the boundaries it is real. As we move along the rectangle of half periods, pu decreases from $+\infty$ to $-\infty$ and is real. u must vary along such a path as will keep k_1 and k_3 real and positive. To find k_1 and k_3 in terms of elliptic functions, we proceed as follows

$$x = 1/(p\tau - pu) = 1/[p\tau - 1/u^2 - c_2u^2 \cdots]$$

and this equals zero for $u = 0$.

Expanding $p(u+\tau)$ in a Taylor's series we obtain for y the following:

$$\begin{aligned} y &= 1/[p\tau - p(u+\tau)] = 1/[p\tau - (p\tau + up'\tau + u^2p''\tau/2! + \cdots)] \\ &= 1/[-up'\tau(1 + up''\tau/2!\,p'\tau + u^2p'''\tau/3!\,p'\tau + \cdots)] \\ &= (-1/up'\tau)[1 - up''\tau/2!\,p'\tau - u^2p'''\tau/3!\,p'\tau + \cdots]. \end{aligned}$$

In a similar way we obtain

$$z = (1/up'\tau)[1 + up''\tau/2!\,p'\tau - u^2p'''\tau/3!\,p'\tau + \cdots];$$

whence it follows

$$x + y + z = p''\tau/p'^2\tau + \text{a function of } u.$$

when $u = 0$, this function vanishes. Hence since $x + y + z = k_1$ is a constant, we have

$$k_1 = p''\tau/p'^2\tau.$$

From 2.6 we have $k_3 = k_1^2/12p\tau = p''^2\tau/12p'^4\tau \cdot p\tau$. Since τ is a third of a period, we have $12p\tau p'^2\tau = p''^2\tau$; for from the identity $2pu + p(2u) = p''^2u/4p'^2u$, we obtain on letting $u = \tau$,

$$2p\tau + p(2\tau) = p''^2\tau/4p'^2\tau.$$

But $p\tau = p(2\tau)$, and hence $12p\tau \cdot p'^2\tau = p''^2\tau$. Putting this back in the expression for k_3 above we obtain

---

* Todhunter's *Algebra*, Page 422.

$$k_3 = 1/p'^2\tau.$$

To keep k_1 and k_3 real and positive, $p'\tau$ must be real, and $p''\tau$ must be real and positive. Both of these conditions hold when τ lies on the real axis. Furthermore the path along which u moves on the cell must keep 2. 7′ positive and real. There is only one choice: u must move along the path which joins the mid-points of the vertical sides of the cell. Thus in our problem $u = \omega_2 + v$, where v is real and varies from 0 to $2\omega_1$.

3. *The Isosceles Cases.* Knowing the path along which u must move, we will next determine when the triangle becomes isosceles. Let us consider first $x = y$. Then $p(u+\tau) = pu$, hence $u + \tau = \pm u + 2m_1\omega_1 + 2m_2\omega_2$. The case of interest here is the one leading to $2u + \tau = 2m_1\omega_1 + 2m_2\omega_2$; but $u = \omega_2 + v$,

$$\therefore \quad 2\omega_2 + 2v + \tau = 2m_1\omega_1 + 2m_2\omega_2;$$

$$\begin{array}{llll} \text{thus} & 2v + \tau = 0, & \text{whence} & v = -\tau/2 = 5\tau/2, \\ \text{or} & = 3\tau, & \text{whence} & v = \tau, \\ \text{or} & = 6\tau, & \text{whence} & v = 5\tau/2, \\ \text{or} & = 9\tau, & \text{whence} & v = 4\tau = \tau. \end{array}$$

Hence the sides a and b of the triangle become equal for two values of u; viz.,

$$u = 5\tau/2 + \omega_2 \quad \text{and} \quad u = \tau + \omega_2.$$

Next we will consider $x = z$. Then $p(u-\tau) = pu$ and going through a similar argument we find the sides a and c of our triangle are equal when

$$u = \omega_2 + \tau/2, \quad \text{and} \quad u = \omega_2 + 2\tau.$$

Finally we have $y = z$ when $p(u+\tau) = p(u-\tau)$. In this case we find that

$$u = \omega_2 \quad \text{and} \quad u = \omega_2 + 3\tau/2,$$

which tells us that our triangle was initially isosceles as $b = c$.

We can sum up the results of this section thus: Starting isosceles, the triangle becomes isosceles at every sixth of a period as u moves along the path $u = \omega_2 + v$, starting with $v = 0$.

4. *The Positional Equations.* We shall determine Λ as a function of the elliptic parameter u. From 1. 1 we have

$$da/d\Lambda = \cos B - \cos C = (-2k_1/abc)(s-a)[(s-c)-(s-b)]$$

$$\therefore \quad da/d\Lambda = (-2k_1k_3/abc)[p'up'\tau/(pu-p\tau)^2].$$

Now $s-a=1/(p\tau-pu)$. Therefore $da/du=-p'u/(p\tau-pu)^2$, and $d\Lambda/du=abc/2k_1k_3^{1/2}$; whence

$$d\Lambda/du=-k_3^{1/2}/2k_1+k_3^{1/2}k_1^2/8k_3-(k_3^{1/2}/2)[pu+p(u+\tau)+p(u-\tau)].$$

Calling $K=-k_3^{1/2}/2k_1+k_3^{1/2}k_1^2/8k_3$, we have

$$d\Lambda=Kdu-(k_3^{1/2}/2)[pu+p(u+\tau)+p(u-\tau)]du.$$

4.1 $$\therefore\quad \Lambda=Ku+(k_3^{1/2}/2)[\zeta u+\zeta(u+\tau)+\zeta(u-\tau)]+c.$$

To determine the constant of integration, let $v=0$, then $u=\omega_2$ and $c=K\omega_2+3k_3^{1/2}\eta_2/2$.

Thus for a particular position of u along its path, we can determine the distance the affices have moved.

As the affices move, the rates of change of the angles θ, ϕ, and ψ are given by a set of equations (see Section 1) which we will call positional equations:

4.2 $$\left\{\begin{aligned} a d\theta/d\Lambda&=\sin B+\sin C,\\ b d\phi/d\Lambda&=\sin C+\sin A,\\ c d\psi/d\Lambda&=\sin A+\sin B.\end{aligned}\right.$$

Expressing these in terms of elliptic functions, we have

$$\begin{aligned} d\theta/d\Lambda&=(\sin B+\sin C)/a\\ &=[2(k_1k_3)^{1/2}/abc]\,[(b+c)/a].\end{aligned}$$

We have already found

$$d\Lambda/du=abc/2k_1k_3^{1/2}$$

Hence
$$\begin{aligned} d\theta/du&=(1/k_1^{1/2})[(b+c)/a]\\ &=(1/k_1^{1/2})\{1+2/[k_1(p\tau-pu)-1]\}.\end{aligned}$$

If we make the substitution

4.3 $$pv_0=p\tau-1/k_1,$$

where pv_0 is a constant, we get $d\theta=(1/k_1^{1/2})[1+2/k_1(pv_0-pu)]du$.*

Multiplying both sides by $p'v_0$ and integrating, we obtain

$$\theta p'v_0=(1/k_1^{1/2})[up'v_0+\frac{2}{k_1}\ (\log\frac{\sigma(u+v_0)}{\sigma(u-v_0)}-2u\zeta v_0)]+C_1.$$

But, by putting for pv_0 its value given in 4.3, in the identity

$$p'v_0=4p^3v_0-g_2pv_0-g_3$$

we find that $p'v_0=2i/k_1^{3/2}$.

* Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 185.

The three angles of the triangle are then given by the following equations:

$$4.4\quad\left\{\begin{aligned} i\theta &= iu/k_1^{1/2} + \log \frac{\sigma(u+v_0)}{\sigma(u-v_0)} - 2u\zeta v_0 + C_1, \\ i\phi &= i(u+\tau)/k_1^{1/2} + \log \frac{\sigma(u+\tau+v_0)}{\sigma(u+\tau-v_0)} - 2(u+\tau)\zeta v_0 + C_2, \\ i\psi &= i(u-\tau)/k_1^{1/2} + \log \frac{\sigma(u-\tau+v_0)}{\sigma(u-\tau-v_0)} - 2(u-\tau)\zeta v_0 + C_3. \end{aligned}\right.$$

The equations 4.4 are important since they tell us the position of the triangle for a particular value of the parameter u.

5. *Introducing the q-series.* The representation of the elliptic functions by the q-series was invented by Jacobi * and is most important for practical problems. His invention made it possible to express doubly periodic functions in an infinite series, the terms of which are singly periodic functions. The problem we are interested in is the study of the paths of the vertices and of the center of gravity of the triangle. First, however, we will express the results already obtained as q-series.

Since we know the network of periods is rectangular, let us choose a rectangle standing upon a smaller side.

Put $2\omega_1 = \pi$ † and then $2\omega_2 = ir\pi$ where $r > 1$. Hence we have $q = e^{i\pi}\omega_2/\omega_1 = e^{-r\pi}$, and the larger we take r the smaller q becomes.

pu expressed as a q-series is ‡

$$pu = -(\eta_1/\omega_1) + (\pi/2\omega_1)^2 1/\sin^2(\pi u/2\omega_1) - 2(\pi/\omega_1)^2 \sum_{n=1}^{\infty} [nq^{2n}/(1-q^{2n})] \cos nu(\pi/\omega_1).$$

For $2\omega_1 = \pi$, we have

$$pu = -2\eta_1/\pi + 1/\sin^2 u - 8 \sum [nq^{2n}/(1-q^{2n})] \cos 2nu;$$

and for $\tau = \pi/3$, we obtain

$$p\tau = -2\eta_1/\pi + 1/\sin^2 \pi/3 - 8 \sum [nq^{2n}/(1-q^{2n})] \cos 2n(\pi/3).$$

we are interested in the values for pu obtained for u moving along a line from ω_2 to $\omega_2 + 2\pi$ or, what is the same thing, for $u = \omega_2 + v$.

---

* Jacobi, *Fundamenta Nova*; Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 425.

† When we have put $2\omega_1 = \pi$, our unit is fixed and we are talking about a particular triangle. To generalize we merely multiply x, y, and z by μ (an arbitrary constant), and the discussion is the same.

‡ Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 426.

$$p(\omega_2 + v) = -2\eta_1/\pi - 8 \sum [nq^n/(1 - q^{2n})] \cos 2nv.$$

Expressed as q-series, we have

5. 1
$$\begin{cases} x = 1/(p\tau - pu) = (3/4)[1 - 6q \cos 2v + q^2(15 + 6 \cos 4 \\ y = 1/[p\tau - p(u + \tau)] = (3/4)[1 - 6q \cos 2(v + \pi/3) \\ \qquad + q^2(15 + 6 \cos 4(v + \pi/3) \\ z = 1/[p\tau - p(u - \tau)] = (3/4)[1 - 6q \cos 2(v - \pi/3) \\ \qquad + q^2(15 + 6 \cos 4(v - \pi/3) \end{cases}$$

5. 2 $$k_1 = x + y + z = (9/4)(1 + q^2 + \cdot\cdot\cdot).$$

5. 3 $$k_3 = xyz = (27/64)(1 + 18q^2 + \cdot\cdot\cdot).$$

5. 4 $$g_2 = 4/3 + 320(q^2 + \cdot\cdot\cdot).\dagger$$

5. 5 $$g_3 = 8/27 - (2^6 \cdot 7/3)(q^2 + \cdot\cdot\cdot).\dagger$$

5. 6 $$\Delta = 2^{12}q^2 + \cdot\cdot\cdot.\dagger$$

We are now prepared to explain our choice of a rectangular ce upon a smaller side. If $q = 0$ we see from 5. 6 that the discrimin But this says $k_1{}^3 = 27k_3$ or that $a = b = c$ which is the equil Once equilateral the triangle stays equilateral, and the vertices circle. It is easy to show that the triangle will never become unless it is that way initially. We shall consider the nearly equil hence we want q to be small. We can also express Λ, θ, ϕ, and ψ We had $d\Lambda/dv = abc/2k_1k_3{}^{1/2}$; hence, $d\Lambda/dv = (2/3^{1/2})(1 + 57q^2/4$

5. 7 $$\therefore \quad \Lambda = (2/3^{1/2})(1 + 57q^2/4 + \cdot\cdot\cdot)v + \Lambda_0,$$

where $\Lambda_0 = 0$. Also

$$d\theta/dv = (1/k_1{}^{1/2})\{1 + 2/[k_1(p\tau - pu) - 1]\},$$

$$d\theta/dv = (2/3)[2 - 9q \cos 2v - 3q^2/2 + 45q^2 \cos 4v/2 + \cdot$$

5. 8 $$\therefore \quad \theta = 4v/3 - 3q \sin 2v - q^2v + 15q^2 \sin 4v/4 + \cdot\cdot\cdot +$$

We can choose our triangle to make $\theta_0 = 0$. In our original when $v = 0$, $u = \omega_2$ and our triangle is isosceles. Let us choose or to be initially parallel to the side a. We must determine the c integration for ϕ and ψ.

$$\phi - k = (4/3)v - 3q \sin 2(v + \pi/3)$$
$$- q^2(v + \pi/3) + (15/4)q^2 \sin 4(v + \pi/3$$
$$\phi_0 - k = -3q \sin(2\pi/3) - q^2 \cdot \pi/3 + (15/4)q^2 \sin (4\pi/$$

whence

5. 9 $$\phi - \phi_0 = (4/3)v + 3q[\sin 2\pi/3 - \sin 2(v + \pi/3)]$$
$$- q^2v + (15/4[\sin 4(v + \pi/3) - \sin (4\pi/3$$

---

* Halphen, *Traité des Fonctions Elliptiques*, Vol. 1, p. 426.

† Harkness and Morley, *Theory of Functions*, pp. 322-324.

In the same way

$$5.10 \qquad \psi - \psi_0 = (4/3)v - 3q[\sin 2\pi/3 + \sin 2(v-\pi/3) \\ - q^2 v + (15/4)[\sin 4(v-\pi/3) + \sin(4\pi/3)] + \cdots.$$

In order to determine ϕ_0 and ψ_0 consider

$$a = k_1 - 1/(p\tau - pu) \\ = 3/2 - (9/2)q\cos 2v + [45/2 - (9/2)\cos 4v]\,q^2 + \cdots.$$

For $v = 0$; $a = a_0$,

$$\therefore \quad a_0 = (3/2)[1 - 3q + 12q^2 + \cdots],$$

and in a similar way, we find

$$b_0 = c_0 = (3/2)[1 + 3q/2 + 33q^2/2 + \cdots].$$

Hence

$$5.11 \qquad \cos\phi_0 = -\frac{a_0}{2b_0} = -\frac{1 - 3q + 12q^2 + \cdots}{2 + 3q + 33q^2 + \cdots}$$

From this we get

$$5.12 \qquad e^{i\phi_0} + e^{-i\phi_0} = -1 + 9q/2 - 9q^2/4 + \cdots.$$

$$\therefore \quad e^{i\phi_0} = \omega + (i3^{3/2}/2)\omega^2 q + \cdots,$$

and

$$i\phi_0 = \log(\omega + (i3^{3/2}/2)\omega^2 q + \cdots.$$

To evaluate ψ_0, we note that it is equal to $-\phi_0$, and hence $\cos\psi_0 = \cos\phi_0$. When we solve 5.12 for $e^{i\phi_0}$, we have two roots resulting from a quadratic equation. They represent the values of $e^{i\phi_0}$ and $e^{i\psi_0}$ respectively. Hence

$$e^{i\psi_0} = \omega^2 - (i3^{3/2}/2)\omega q + \cdots,$$

and

$$i\psi_0 = \log[\omega^2 - i3^{3/2}/2)\omega q + \cdots].$$

Formula 5.11 is very important in that it fixes the value of q once the initial lengths of the sides of the triangle are given.

6. *The Paths of the Vertices and Centroid.* First consider the path of A. $e^{i\theta}$ is the turn from the base line to the side a. A is moving along some path with a direction $e^{(\pi+\theta)i} = -e^{i\theta}$. Since this equals $dA/d\Lambda$, we have

$$dA/d\Lambda = \exp\{i[4v/3 - 3q\sin 2v - q^2 v + 15q^2\sin 4v/2 + \cdots]\} \\ = -\exp(4iv/3)\{1 - 3iq\sin 2v \\ + q^2[-(9/4) - iv + (9/4)\cos 4v + (15i/4)\sin 4v] + \cdots\}, \quad \text{and}$$

$$d\Lambda/dv = 2(1 + 57q^2/4 + \cdots)/3^{1/2}.$$

$$\therefore \quad dA/dv = -(2/3^{1/2})\exp[i(4/3)v]\,[1 - 3iq\sin 2v + q^2(12 - iv + 9\cos 4v/4 + 15i\sin 4v/4) + \cdots]$$
$$= (-2/3^{1/2})\exp[i(4/3)v]\{1 - 3q[\exp(2iv) - \exp(-2iv)]/2 + q^2[48 - 4iv + 12\exp(4iv) - 3\exp(-4iv)]/4 + \cdots\}.$$

Put $\exp(2iv) = t^3$, then $dv = 3dt/2it$, and

$$dA/dt = i3^{1/2}[t - 3q(t^4 - t^{-2})/2 + (q^2/4)(48t - 6t\log t + 12t^7 - 3t^{-5}) + \cdots].$$

6.1 $\therefore \quad A = A_0 + i3^{1/2}\{t^2/2 - (3q/10t)(t^6 + 5) + (q^2/16)[102t^2 - 12t^2\log t + 6t^8 + 3t^{-4}] + \cdots\}.$

We can determine A_0 by taking $A = 0$ when $t = 1$. This represents the path along which A moves. The logarithmic term tells us the path is not closed but continually shifts over the plane.

By a similar method the equations of the paths of B and C, expanded as far as the first degree term in q, are found to be

6.2 $$B = B_0 + i3^{1/2}\left(\frac{\omega t^2}{2} - \frac{i3^{3/2}qt^2}{4} - \frac{3\omega q(\omega t^6 + 5\omega^2)}{10t} + \cdots\right),$$

6.3 $$C = C_0 + i3^{1/2}\left(\frac{\omega^2 t^2}{2} + \frac{i3^{3/2}qt^2}{4} - \frac{3\omega^2 q(\omega^2 t^6 + 5\omega)}{10t} + \cdots\right)$$

where B_0 and C_0 are determined in the same manner as A_0.

To obtain the path of the centroid g we have, noting that $3g = A + B + C$,

that $\quad g = (A_0 + B_0 + C_0)/3 + 3^{3/2}q/2it + \cdots,$

where $\quad A_0 + B_0 + C_0 = 3^{3/2}iq/2 + \cdots.$

Hence $\quad g = (3^{3/2}q/2i)[1/(t-1)] + \cdots.$

Periodic Orbits in the Problem of Three Bodies with Repulsive and Attractive Forces.

By Daniel Buchanan.

1. *Introduction.* This paper deals with periodic orbits described by two mutually repellant infinitesimal bodies which are attracted by a finite body. The forces of repulsion and attraction are assumed to vary according to the Newtonian law of the inverse square. Two types of periodic orbits for this system were obtained by Rawles.* In the first type, which will be here designated as the *circular orbits,* the repellant particles move in equal circles the planes of which are parallel. The line joining the centres of these circles is normal to their planes and is bisected by the centre of gravity of the finite body. The particles remain on the same generating line of the cylinder through these circles.

In the orbits of the second type, here designated as the *arc orbits,* the three bodies remain in the same plane. The infinitesimal bodies oscillate in arcs of curves, which are symmetrically situated with respect to the finite body. Langmuir † first calculated these orbits by numerical integration and they are also discussed by Van Vleck.‡

The problem considered in the present paper deals with periodic oscillations in the vicinity of the circular orbits. Only the construction of these orbits is made but the convergence of the solutions obtained is assured by a theorem due to MacMillan.§ The author begs to acknowledge the assistance of Mr. H. D. Smith, M. A.,¶ in checking certain algebraic expressions in the construction and in making the computation for the numerical examples.

Second genus orbits in the vicinity of the arc orbits have also been obtained by the author but they are discussed in another article.‖

* Rawles, "Two Classes of Periodic Orbits with Repelling Forces," *Bulletin of the American Mathematical Society,* Vol. 34, No. 5 (1928), pp. 618-630.

† Langmuir, *Physical Review,* Vol. 17 (1921), pp. 339-353.

‡ Van Vleck, "Quantum Principles and Line Spectra," *Bulletin of the National Research Council,* Vol. 10, Part 4, No. 54, p. 89.

§ MacMillan, *Transactions of the American Mathematical Society,* Vol. 13, No. 2, pp. 146-158.

¶ Smith, A thesis submitted in the Department of Mathematics for the degree of M. A. in the University of British Columbia.

‖ Buchanan, "Second Genus Orbits for the Helium Atom," *Transactions of the Royal Society of Canada,* Third Series, Vol. 23, Sec. 3 (1929), pp. 227-245.

As there is a similarity between the three bodies in this problem and the helium atom, we shall refer to the finite body as the nucleus and to the particles as electrons. No use, however, is made of the quantum mechanics nor of Larmor's theorem.*

2. *The Circular Orbits.* The units of time and space will be chosen so that the gravitational constant of attraction is unity. Let k^2 denote the ratio of the repulsion to the attraction. Then the force function of the system is

$$U = 1/\rho_1 + 1/\rho_2 - k^2/\Delta,$$

where ρ_1 and ρ_2 are the distances between the electrons and the nucleus, and Δ is the distance between the electrons. If we take a system of rectangular coördinates with the origin at the nucleus and denote the coördinates of the electrons as (x_j, y_j, z_j), $(j = 1, 2)$, then the differential equations defining their motion are

$$\begin{aligned} &x_j'' = \partial U/\partial x_j, \quad y_j'' = \partial U/\partial y_j, \quad z_j'' = \partial U/\partial z_j, \\ &\rho_j^2 = x_j^2 + y_j^2 + z_j^2, \qquad\qquad (j = 1, 2), \\ &\Delta^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2. \end{aligned} \tag{1}$$

When the restrictions

$$x_1 = - x_2, \quad y_1 = y_2, \quad z_1 = z_2 \tag{2}$$

are made, as in Rawles' paper, the differential equations become

$$\begin{aligned} x'' &= - x/\rho^3 + k^2/4x^2, \\ y'' &= - y/\rho^3, \\ z'' &= - z/\rho^3, \end{aligned} \tag{3}$$

where the subscripts 1 or 2 have been dropped. These equations possess the integrals

$$\begin{aligned} \tfrac{1}{2}(x'^2 + y'^2 + z'^2) &= 1/\rho - k^2/4x + \text{const.}, \\ y'z - yz' &= \text{const.} \end{aligned} \tag{4}$$

The solutions of the differential equations are

$$\begin{aligned} x &= (k^2/4)^{1/3} = m, \text{ say}, \\ y &= (1 - m^2)^{1/2} \sin (t - t_0), \\ z &= (1 - m^2)^{1/2} \cos (t - t_0), \end{aligned} \tag{5}$$

which are the circular solutions obtained by Rawles. They denote the circles

* Larmor, *Philosophical Magazine*, V, Vol. 44 (1897), p. 503; Richardson, *The Electron Theory of Matter* (1916), p. 258.

with centres at $(\pm m, 0, 0)$, radii $(1-m^2)^{1/2}$ and whose planes are parallel to the yz-plane. The electrons rotate in these orbits from the positive z-axis to the positive y-axis. If the solutions are to be real, m^2 cannot exceed unity. When $m^2 = 1$, however, the solutions reduce to point circles but this simple case will be excluded from our consideration.

We shall refer only to the one circle, viz., that having its centre at $(m, 0, 0)$.

Orbits of Three Dimensions.

3. *The Differential Equations.* Let the motion be referred to a system of rotating axes x, η, ξ. The x-axis remains unchanged while the $\eta\xi$-axes rotate in the yz-plane in the direction in which the electrons move and with their angular velocity. Further, let $\eta = -y$, $\xi = z$ at $t = t_0$. Then the necessary transformations are

$$(6) \qquad \begin{aligned} y &= -\eta \cos(t-t_0) + \xi \sin(t-t_0), \\ z &= \eta \sin(t-t_0) + \xi \cos(t-t_0), \end{aligned}$$

and the differential equations of motion (3) become

$$(7) \qquad \begin{aligned} x'' &= -x/\rho^3 + k^2/4x^2, \\ \eta'' + 2\xi' - \eta &= -\eta/\rho^3, \\ \xi'' - 2\eta - \xi &= -\xi/\rho^3. \end{aligned}$$

A particular solution of these equations is

$$(8) \qquad x = m, \quad \eta = 0, \quad \xi = (1-m^2)^{1/2},$$

which are the equations of the circular orbit with respect to the rotating axes.

In order to determine deviations from the circular orbit, let

$$(9) \qquad \begin{aligned} x &= m + \gamma p, \\ \eta &= 0 + \gamma q, \\ \xi &= (1-m^2)^{1/2} + \gamma r, \\ t - t_0 &= (1+\delta)^{1/2}\tau, \end{aligned}$$

where

p, q, r are new dependent variables,
γ is a parameter representing the scale factor of the new orbits,
δ is a constant depending upon γ,
τ is the new independent variable.

When equations (9) are substituted in (7) and the factor γ is divided out,

the following differential equations are found, the dots denoting derivation with respect to τ;

$$
\begin{aligned}
&\ddot{p} + 3(1+\delta)(1-m^2)p - 3(1+\delta)m(1-m^2)^{1/2}r \\
&\qquad = (1+\delta)[\gamma P_2 + \gamma^2 P_3 + \cdots + \gamma^j P_{j+1} + \cdots], \\
(10)\quad &\ddot{q} + 2(1+\delta)^{1/2}\dot{r} = (1+\delta)[\gamma Q_2 + \cdots + \gamma^j Q_{j+1} + \cdots], \\
&\ddot{r} - 2(1+\delta)^{1/2}\dot{q} - 3(1+\delta)(1-m^2)r - 3(1+\delta)m(1-m^2)^{1/2}p \\
&\qquad = (1+\delta)[\gamma R_2 + \cdots + \gamma^j R_{j+1} + \cdots],
\end{aligned}
$$

where P_j, Q_j, R_j $(j = 2, 3, \cdots)$ are polynomials in p, q, r of degree j. In P_j and R_j, q enters to even degrees only, while in Q_j it enters to odd degrees only. So far as the computation has been carried out we have

$$
\begin{aligned}
P_2 &= 3(1/m + 3m/2 - 5m^3/2)p^2 + 3mq^2/2 \\
&\quad - 3m(2 - 5m^2/2)r^2 + 3(1-m^2)^{1/2}(1-5m^2)pr, \\
P_3 &= (3/2 - 15m^2 + 35m^4/2)p^3 + (15m/2)(1-m^2)^{1/2}(7m^2-3)p^2r \\
&\quad + (3/2)(1-5m^2)pq^2 - 3(2 - 35m^2/2 + 35m^4/2)pr^2 \\
&\quad - (15m/2)(1-m^2)^{1/2}q^2r + 5(2 - 7m^2/2)r^3, \\
Q_2 &= 3mpq + 3(1-m^2)^{1/2}qr, \\
Q_3 &= (3/2)(1-5m^2)p^2q - 15m(1-m^2)^{1/2}pqr + 3q^3/2 \\
&\quad - 3(2 - 5m^2/2)r^2q, \\
R_2 &= (3/2)(1-m^2)^{1/2}(1-5m^2)p^2 - 3m(4-5m^2)pr \\
&\quad + (3/2)(1-m^2)^{1/2}q^2 - 3(1-m^2)^{1/2}(1-5m^2/2)r^2, \\
R_3 &= (5m/2)(1-m^2)^{1/2}(7m^2-3)p^3 \\
&\quad + 3[1/2 + 15m^2 - 35m^4/2 - (5m/2)(1-m^2)^{1/2}]p^2r \\
&\quad - (15m/2)(1-m^2)^{1/2}pq^2 + 15m(1-m^2)^{1/2}(2 - 7m^2/2)pr^2 \\
&\quad + 3[1/2 - 5m(1-m^2)^{1/2}]q^2r \\
&\quad - [27/2 - 15m^2 - (35/2)(1-m^2)^{5/2}]r^3.
\end{aligned}
$$

On integrating (10, b) we obtain

$$
\begin{aligned}
(11)\qquad \dot{q} &= -2(1+\delta)^{1/2}r + C \\
&\quad + (1+\delta)\int (\gamma Q_2 + \cdots + \gamma^j Q_{j+1} \cdots)\, d\tau
\end{aligned}
$$

where C is the constant of integration. As q and r are later developed as power series in γ we shall put

$$
(12)\qquad C = C_1^{(0)} + C_1^{(1)}\gamma + \cdots + C_1^{(n)}\gamma^n + \cdots.
$$

When the substitutions are made for C in (11) and for $\dot{q}$ in (10, c) we obtain, on repeating (10, a) and (11) for reference,

$$
\begin{aligned}
&\ddot{p} + 3(1+\delta)(1-m^2)p - 3(1+\delta)m(1-m^2)^{1/2}r \\
&\qquad = (1+\delta) \sum_{j=1}^{\infty} \gamma^j P_{j+1}, \\
(13)\qquad &\dot{q} = -2(1+\delta)^{1/2}r + (1+\delta)\int \sum_{j=1}^{\infty} \gamma^j Q_{j+1} d\tau + \sum_{j=0}^{\infty} C_1^{(j)}\gamma^j, \\
&\ddot{r} + (1+\delta)(1+3m^2)r - 3(1+\delta)m(1-m^2)^{1/2}p \\
&\qquad = (1+\delta) \sum_{j=1}^{\infty} \gamma^j R_{j+1} + 2(1+\delta)^{1/2} \sum_{j=0}^{\infty} C_1^{(j)}\gamma^j \\
&\qquad + 2(1+\delta)^{3/2} \int \sum_{j=1}^{\infty} \gamma^j Q_{j+1} d\tau.
\end{aligned}
$$

We shall now take (13) as the three defining equations for p, q, r.

4. *The Equations of Variation and their Solutions.* If we consider only the terms of the equations (13) which are independent of γ we obtain the equations of variation. They are

$$
(14)\qquad
\begin{aligned}
&\ddot{p} + 3(1-m^2)p - 3m(1-m^2)^{1/2}r = 0, \\
&\dot{q} + 2r = C_1^{(0)}, \\
&\ddot{r} + (1+3m^2)r - 3m(1-m^2)^{1/2}p = 2C_1^{(0)}.
\end{aligned}
$$

The first and third equations of (14) are independent of the second and will be considered first. We shall make use of the operator D to denote $d/d\tau$. Then (14, a) and (14, c) may be expressed as

$$
(15)\qquad
\begin{aligned}
&[D^2 + 3(1-m^2)]\,p - 3m(1-m^2)^{1/2}r = 0, \\
&-3m(1-m^2)^{1/2}p + [D^2 + 1 + 3m^2]r = 2C_1^{(0)}.
\end{aligned}
$$

The functional determinant of these equations is

$$
(16)\qquad \mathfrak{D} = \begin{vmatrix} D^2 + 3(1-m^2), & -3m(1-m^2)^{1/2} \\ -3m(1-m^2)^{1/2}, & D^2 + 1 + 3m^2 \end{vmatrix} = D^4 + 4D^2 + 3(1-m^2).
$$

On equating $\mathfrak{D}$ to zero, as in the method of solving sets of linear differential equations with constant coefficients, we find the roots

$$
D^2 = -2 + (1+3m^2)^{1/2}, \quad -2 - (1+3m^2)^{1/2}.
$$

As m^2 must be less than 1 in order that the circular solutions shall be real, both roots for D^2 are therefore negative. If we put

$$
-2 + (1+3m^2)^{1/2} = -\sigma_1^2, \quad -2 - (1+3m^2)^{1/2} = -\sigma_2^2,
$$

then

$$
D = \pm i\sigma_1, \quad \pm i\sigma_2,
$$

and the complementary functions of (15) are thus found to be

$$
\begin{aligned}
p &= A_1e^{i\sigma_1\tau} + A_2e^{-i_1\sigma\tau} + A_3e^{i\sigma_2\tau} + A_4e^{-i\sigma_2\tau},\\
r &= B_1e^{i\sigma_1\tau} + B_2e^{-i\sigma_1\tau} + B_3e^{i\sigma_2\tau} + B_4e^{-i\sigma_2\tau},
\end{aligned}
\tag{17}
$$

where A_j, B_j, $(j=1,\cdots,4)$ are constants of integration. Only four of these constants are independent as the following relations hold,

$$
\begin{aligned}
A_j &= \omega_\nu B_j, \qquad\qquad (j=1,2,3,4;\ \nu=1,2),\\
\omega_1 &= 3m(1-m^2)^{1/2}/[1-3m^2+(1+3m^2)^{1/2}],\\
\omega_2 &= 3m(1-m^2)^{1/2}/[1-3m^2-(1+3m^2)^{1/2}].
\end{aligned}
\tag{18}
$$

There are therefore three sets of generating solutions, viz.,

I
$$
\begin{aligned}
p &= \omega_1(B_1e^{i\sigma_1\tau} + B_2e^{-i\sigma_1\tau}),\\
r &= B_1e^{i\sigma_1\tau} + B_2e^{-i\sigma_1\tau};\\
&\text{Period} = P_1 = 2\pi/\sigma_1.
\end{aligned}
$$

II
$$
\begin{aligned}
p &= \omega_2(B_3e^{i\sigma_2\tau} + B_4e^{-i\sigma_2\tau}),\\
r &= B_3e^{i\sigma_2\tau} + B_4e^{-i\sigma_2\tau};\\
&\text{Period} = P_2 = 2\pi/\sigma_2.
\end{aligned}
$$

III
$$
\begin{aligned}
p &= \omega_1(B_1e^{i\sigma_1\tau} + B_2e^{-i\sigma_1\tau}) + \omega_2(B_3e^{i\sigma_2\tau} + B_4e^{-i\sigma_2\tau}),\\
r &= B_1e^{i\sigma_1\tau} + B_2e^{-i\sigma_1\tau} + B_3e^{i\sigma_2\tau} + B_4e^{-i\sigma_2\tau};\\
&\text{Period} = P_3 = n_2P_1 = n_1P_2.
\end{aligned}
$$

The last solutions, III, exist only when σ_1 and σ_2 are commensurable, i. e., when

$$\sigma_1/\sigma_2 = n_1/n_2,$$

where n_1 and n_2 are relatively prime integers.

Orbits are constructed in the sequel by using only the first two generating solutions. The construction of orbits having generating solutions III was attempted but abandoned on account of the complexity of the problem.

5. *Outline of the Construction of Periodic Solutions.* There is the same construction for orbits having the generating solutions I or II except for the subscripts 1 and 2, respectively, on σ and ω. We shall therefore drop these subscripts and restore them in the final solutions.

We propose to show that p, q, r, δ can be determined as power series in γ so that p, q, r shall be periodic with the period $P(=P_1$ or $P_2)$ and shall satisfy certain initial conditions, to be discussed presently. Accordingly we put

$$
\begin{aligned}
p &= \sum_{j=0}^{\infty} p_j\gamma^j, \qquad q = \sum_{j=0}^{\infty} q_j\gamma^j,\\
r &= \sum_{j=0}^{\infty} r_jq^j, \qquad \delta = \sum_{j=1}^{\infty} \delta_j\gamma^j.
\end{aligned}
\tag{19}
$$

Let these substitutions be made in (13) and let the resulting equations be cited as (13′). On equating the coefficients of the various powers of γ in (13′) we obtain sets of differential equations in p_j, q_j, r_j. We propose to show that these equations can be integrated and that the various δ_j and the constants of integration at each step can be determined so that p_j, q_j, r_j shall be periodic and shall satisfy the initial conditions, now to be discussed.

6. *The Initial Conditions.* It will be observed in the next section that at each step of the integration four arbitrary constants arise which are not determined by the periodicity conditions. We therefore impose four initial conditions. Let us suppose that

$$\dot{p}(0) = \dot{r}(0) = q(0) = 0, \qquad r(0) \neq 0.$$

As r carries the factor γ in (9) we may take $r(0) = 1$ without loss of generality. When these initial conditions are imposed upon (19) we obtain

$$(20) \qquad \begin{aligned} &\dot{p}_j(0) = \dot{r}_j(0) = q_j(0) = 0, && (j = 0, 1, 2 \cdots), \\ &r_1(0) = 1, \quad r_j(0) = 0, && (j = 2, 3, 4, \cdots). \end{aligned}$$

7. *Construction of the Solutions.*

Terms independent of γ. When we equate the coefficients of the terms in (13′) which are independent of γ we obtain equations which are the same as (14) except for the subscript 0 on p, q and r. The solutions which have the period P_1 or P_2, except for certain terms in τ, are

$$(21) \qquad \begin{aligned} p_0 &= \omega(B_1^{(0)} e^{i\sigma\tau} + B_2^{(0)} e^{-i\sigma\tau}) + 2m(1 - m^2)^{-1/2} C_1^{(0)}, \\ q_0 &= (2i/\sigma)(B_1^{(0)} e^{i\sigma\tau} - B_2^{(0)} e^{-i\sigma\tau}) - 3C_1^{(0)}\tau + C_2^{(0)}, \\ r_0 &= B_1^{(0)} e^{i\sigma\tau} + B_2^{(0)} e^{-i\sigma\tau} + 2C_1^{(0)}, \end{aligned}$$

where B and C, here and henceforth, with various subscripts and superscripts are constants of integration.

In order to satisfy the periodicity conditions we must put $C_1^{(0)} = 0$. When we impose the condition $\dot{p}_0(0) = 0$ we obtain $B_1^{(0)} = B_2^{(0)}$, and consequently the condition $\dot{r}_0(0) = 0$ is satisfied. Then from $q_0(0) = 0$ we obtain $C_2^{(0)} = 0$ and from $r_0(0) = 1$ we have

$$B_1^{(0)} = B_2^{(0)} = 1/2.$$

The periodic solutions at this step which satisfy the initial conditions then become

$$(22) \qquad p_0 = \omega \cos \sigma\tau, \quad q_0 = -(2/\sigma) \sin \sigma\tau, \quad r_0 = \cos \sigma\tau.$$

Terms in γ. The differential equations arising from the terms in γ in (13′) are

$$
(23)\qquad
\begin{aligned}
&[D^2+3(1-m^2)]p_1-3m(1-m^2)^{1/2}r_1=P^{(1)},\\
&-3m((1-m^2)^{1/2}p_1+[D^2+1+3m^2]r_1=R^{(1)}+2C_1^{(1)},\\
&\dot{q}_1=-2r_1+C_1^{(1)}+\int Q^{(1)}d\tau,
\end{aligned}
$$

where

$$
\begin{aligned}
P^{(1)}&=a_0^{(1)}+\delta_1 a_1^{(1)}\cos\sigma\tau+a_2^{(1)}\cos 2\sigma\tau,\\
Q^{(1)}&=\qquad\quad \delta_1 b_1^{(1)}\sin\sigma\tau+b_2^{(1)}\sin 2\sigma\tau,\\
R^{(1)}&=c_0^{(1)}+\delta_1 c_1^{(1)}\cos\sigma\tau+c_2^{(1)}\cos 2\sigma\tau;\\
a_0^{(1)}&=(3/2)(1/m+3m/2-5m^3/2)\omega^2+(3/2)(1-m^2)^{1/2}(1-5m^2)\\
&\qquad-3m(1-1/\sigma^2-5m^2/4),\\
a_1^{(1)}&=-3(1-m^2)\omega+3m(1-m^2)^{1/2},\\
a_2^{(1)}&=(3/2)(1/m-3m/2-5m^3/2)\omega^2+(3/2)(1-m^2)^{1/2}(1-5m^2)\\
&\qquad-3m(1+1/\sigma^2-5m^2/4),\\
b_1^{(1)}&=1,\quad b_2^{(1)}=-(3/\sigma)[m\omega+(1-m^2)^{1/2}],\\
c_0^{(1)}&=(3/4)(1-m^2)^{1/2}(1-5m^2)\omega^2-3m(2-5m^2/2)\omega\\
&\qquad-3(1-m^2)^{1/2}(1/2-1/\sigma^2-5m^2/4),\\
c_1^{(1)}&=3m(1-m^2)^{1/2}+(1-3m^2),\\
c_2^{(1)}&=(3/4)(1-m^2)^{1/2}(1-5m^2)\omega^2-3m(2-5m^2/2)\omega\\
&\qquad-(3/2)(1-m^2)^{1/2}(1+1/\sigma^2-5m^2/2).
\end{aligned}
$$

The solutions of (23, a and b) will be considered first as (23, c) depends upon r_1. The complementary functions of (23 a) and (23 b) are

$$
(24)\qquad
\begin{aligned}
p_1&=\omega(B_1^{(1)}e^{i\sigma\tau}+B_2^{(1)}e^{-i\sigma\tau})+2m(1-m^2)^{-1/2}C_1^{(1)},\\
r_1&=B_1^{(1)}e^{i\sigma\tau}+B_2^{(1)}e^{-i\sigma\tau}+2C_1^{(1)}.
\end{aligned}
$$

The particular integrals of p_1 and r_1, expressed symbolically, are

$$
(25)\qquad
\begin{aligned}
p_1&=\frac{[D^2+1+3m^2]P^{(1)}+3m(1-m^2)^{1/2}R^{(1)}}{D^4+4D^2+3(1-m^2)},\\
r_1&=\frac{3m(1-m^2)^{1/2}P^{(1)}+[D^2+3(1-m^2)]R^{(1)}}{D^4+4D^2+3(1-m^2)}.
\end{aligned}
$$

In order that p_1 and r_1 shall be periodic the coefficients of $\cos\sigma\tau$ in the numerators of the above expressions must vanish, inasmuch as $-\sigma^2$ is a root of the denominators. Hence

$$
(26)\qquad
\begin{aligned}
&\delta_1[a_1^{(1)}\{-\sigma^2+1+3m^2\}+c_1^{(1)}\{3m(1-m^2)^{1/2}\}]=0,\\
&\delta_1[a_1^{(1)}\{3m(1-m^2)^{1/2}\}+c^{(1)}\{-\sigma^2+3(1-m^2)\}]=0.
\end{aligned}
$$

The functional determinant of $\delta_1 a_1^{(1)}$ and $\delta_1 c_1^{(1)}$ in the above equations is

$$
\sigma^4-4\sigma^2+3(1-m^2),
$$

and this vanishes as $-\sigma^2$ is a root of $\mathfrak{D}$ in (16). Therefore the two equations in (26) are equivalent. They are satisfied only by $\delta_1 = 0$. The particular integrals then become

$$(27)\qquad \begin{aligned} p_1 &= \alpha_0^{(1)} + \alpha_2^{(1)} \cos 2\sigma\tau, \\ r_1 &= \gamma_0^{(1)} + \gamma_2^{(1)} \cos 2\sigma\tau, \end{aligned}$$

where

$$\begin{aligned} \alpha_0^{(1)} &= \frac{1+3m^2}{3(1-m^2)}\, a_0^{(1)} + \frac{m}{(1-m^2)^{1/2}}\, c_0^{(1)}, \\ \alpha_2^{(1)} &= \frac{(1-4\sigma^2+3m^2)a_2^{(1)} + 3m(1-m^2)^{1/2}c_2^{(1)}}{16\sigma^4-16\sigma^2+3(1-m^2)}, \\ \gamma_0^{(1)} &= \frac{m}{(1-m^2)^{1/2}}\, a_0^{(1)} + c_0^{(1)}, \\ \gamma_2^{(1)} &= \frac{3m(1-m^2)^{1/2}a_2^{(1)} - \{4\sigma^2-3(1-m^2)\}c_2^{(1)}}{16\sigma^4-16\sigma^2+3(1-m^2)}. \end{aligned}$$

When (24) and (27) are combined we obtain the complete solutions for p_1 and r_1.

The third equation of (23) can now be integrated, the integral being

$$\begin{aligned} q_1 = (2i/\sigma)\,(B_1^{(1)}e^{i\sigma\tau} - B_2^{(1)}e^{-i\sigma\tau}) - (3c_1^{(1)} + 2\gamma_0^{(1)})\tau \\ + \beta_2^{(1)} \sin 2\sigma\tau + C_2^{(1)}, \end{aligned}$$

where

$$\beta_2^{(1)} = (3/4\sigma^3)\,[m\omega + (1-m^2)^{1/2} - 2\gamma_2^{(1)}].$$

On applying the periodicity and initial conditions to the complete solutions for p_1, q_1, r_1 we obtain

$$\begin{aligned} C_1^{(1)} &= -(2/3)\gamma_0^{(1)}, C_2^{(1)} = 0, \\ B_1^{(1)} &= B_2^{(1)} = (1/6)\gamma_0^{(1)} - (1/2)\gamma_2^{(1)}. \end{aligned}$$

The desired solutions at this step are thus found to be

$$(28)\qquad \begin{aligned} p_1 &= F_0^{(1)} + F_1^{(1)} \cos \sigma\tau + F_2^{(1)} \cos 2\sigma\tau, \\ q_1 &= G_1^{(1)} \sin \sigma\tau + G_2^{(1)} \sin 2\sigma\tau, \\ r_1 &= H_0^{(1)} + H_1^{(1)} \cos \sigma\tau + H_2^{(1)} \cos 2\sigma\tau, \end{aligned}$$

where

$$\begin{aligned} &F_0^{(1)} = \alpha_0^{(1)} - (4m/3)\,(1-m^2)^{-1/2}\gamma_0^{(1)}, \\ &F_1^{(1)} = 2B_1^{(1)}\omega,\ F_2^{(1)} = \alpha_2^{(1)}, \\ &G_1^{(1)} = -(4/\sigma)B_1^{(1)},\ G_2^{(1)} = \beta_2^{(1)}, \\ &H_0^{(1)} = -(1/3)\gamma_0^{(1)},\quad H_1^{(1)} = 2B_1^{(1)},\ H_2^{(1)} = \gamma_2^{(1)}. \end{aligned}$$

Terms in γ^2. It will be necessary to consider the terms in γ^2 in (13′) before the induction to the general term can be made. These terms are

$$
(29)\qquad
\begin{aligned}
&[D^2+3(1-m^2)]\,p_2-3m(1-m^2)^{1/2}r_2=P^{(2)},\\
&-3m(1-m^2)^{1/2}p_2+(D^2+1+3m^2)r_2=R^{(2)}+2C_1^{(2)},\\
&\dot{q}_2=-2r_2+\int Q^{(2)}d\tau-\delta_2 r_0+C_1^{(2)},
\end{aligned}
$$

where

$$
\begin{aligned}
P^{(2)}&=a_0^{(2)}+(\delta_2 d_1^{(2)}+a_1^{(2)})\cos\sigma\tau\\
&\qquad+a_2^{(2)}\cos 2\sigma\tau+a_3^{(2)}\cos 3\sigma\tau,\\
R^{(2)}&=c_0^{(2)}+(\delta_2 d_2^{(2)}+c_1^{(2)})\cos\sigma\tau\\
&\qquad+c_2^{(2)}\cos 2\sigma\tau+c_3^{(2)}\cos 3\sigma\tau,\\
Q^{(2)}&=b_1^{(2)}\sin\sigma\tau+b_2^{(2)}\sin 2\sigma\tau+b_3^{(2)}\sin 3\sigma\tau,\\
d_1^{(2)}&=3m(1-m^2)^{1/2}-3(1-m^2)\omega,\\
d_2^{(2)}&=1-3m^2+3m(1-m^2)^{1/2}\omega.
\end{aligned}
$$

The values of the various a's, b's, and c's were computed by Mr. Smith but his results are omitted here.

The complementary functions and the particular integrals of the first two equations of (29) are the same as (24) and (25), respectively, with the appropriate changes in subscripts and superscripts. The equations similar to (26) which must be satisfied in order that the particular integrals for p_1 and r_1 shall be periodic, are

$$
(30)\qquad
\begin{aligned}
&(1-\sigma^2+3m^2)[\delta_2 d_1^{(2)}+a_1^{(2)}]+3m(1-m^2)^{1/2}[\delta_2 d_2^{(2)}+c_1^{(2)}]=0,\\
&3m(1-m^2)^{1/2}[\delta_2 d_1^{(2)}+a_1^{(2)}]+\{-\sigma^2+3(1-m^2)\}[\delta_2 d_2^{(2)}+c_1^{(2)}]=0.
\end{aligned}
$$

The determinant of the coefficients of the expressions in the brackets [ ] is the same here as in (26) and therefore vanishes. Hence the above equations are identical and can be satisfied by a proper choice of the single arbitrary δ_2. The required value of δ_2 is

$$
(31)\qquad \delta_2=\frac{(\sigma^2-1-3m^2)a_1^{(2)}-3m(1-m^2)^{1/2}c_1^{(2)}}{(1-\sigma^2+3m^2)d_1^{(2)}+3m(1-m^2)^{1/2}d_2^{(2)}}.
$$

When δ_2 is thus determined, the complete solutions for p_2 and r_2 will be periodic and will have the form

$$
(32)\qquad
\begin{aligned}
p_2&=\omega(B_1^{(2)}e^{i\sigma\tau}+B_2^{(2)}e^{-i\sigma\tau})+2m(1-m^2)^{-1/2}C_1^{(2)}\\
&\qquad+\alpha_0^{(2)}+\alpha_2^{(2)}\cos 2\sigma\tau+\alpha_3^{(2)}\cos 3\sigma\tau,\\
r_2&=B_1^{(2)}e^{i\sigma\tau}+B_2^{(2)}e^{-i\sigma\tau}+2C_1^{(2)}\\
&\qquad+\gamma_0^{(2)}+\gamma_2^{(2)}\cos 2\sigma\tau+\gamma_3^{(2)}\cos 3\sigma\tau,
\end{aligned}
$$

where the α's and γ's are linear in the a's and c's.

On substituting (32) in (29 c) and integrating we obtain

$$
\begin{aligned}
q_2&=(2i/\sigma)(B_1^{(2)}e^{i\sigma\tau}-B_2^{(2)}e^{-i\sigma\tau})+(3C_1^{(2)}+2\gamma_0^{(2)})\tau\\
&\quad+C_2^{(2)}+\beta_1^{(2)}\sin\sigma\tau+\beta_2^{(2)}\sin 2\sigma\tau+\beta_3^{(2)}\sin 3\sigma\tau,
\end{aligned}
$$

where

$$\beta_1^{(2)} = -(1/\sigma)\delta_2 - (1/\sigma^2)b_1^{(2)},$$
$$\beta_2^{(2)} = -(1/\sigma)\gamma_2^{(2)} - (1/4\sigma^2)b_2^{(2)},$$
$$\beta_3^{(2)} = -(2/3\sigma)\gamma_3^{(2)} - (1/9\sigma^2)b_3^{(2)}.$$

When the periodicity and initial conditions are applied we have

$$C_1^{(2)} = -(2/3)\gamma_0^{(2)}, \quad C_2^{(2)} = 0,$$
$$B_1^{(2)} = B_2^{(2)} = (1/6)\gamma_0^{(2)} - \gamma_2^{(2)} + \gamma_3^{(2)}.$$

The solutions at the third step are therefore

$$p_2 = \sum_{\nu=0}^{3} F_\nu^{(2)} \cos \nu\sigma\tau,$$
$$q_2 = \sum_{\nu=1}^{3} G_\nu^{(3)} \sin \nu\sigma\tau,$$
$$r_2 = \sum_{\nu=0}^{3} H_\nu^{(2)} \cos \nu\sigma\tau,$$

where

$$F_0^{(2)} = 2m(1-m^2)^{-\frac{1}{2}}c_1^{(2)} + \alpha_0^{(2)},$$
$$F_1^{(2)} = 2\omega B_1^{(2)} \quad F_j^{(2)} = \alpha_j^{(2)}, \qquad (j = 2, 3,),$$
$$G_1^{(2)} = -(4/\sigma)B_1^{(2)} + \beta_1^{(2)}, G_j^{(2)} = \beta_j^{(2)}, \qquad (j = 2, 3),$$
$$H_0^{(2)} = 2c_1^{(2)} + \gamma_0^{(2)},$$
$$H_1^{(2)} = 2B_1^{(2)}, H_j^{(2)} = \gamma_j^{(2)}, \qquad (j = 2, 3).$$

8. *Induction to the General Term.* Let us suppose that the p_j, q_j, r_j have all been determined for $j = 0, \cdots, n-1$ and that they are of the form

$$p_j = \sum_{\nu=0}^{j+1} F_\nu^{(j)} \cos \nu\sigma\tau,$$
$$(33) \qquad q_j = \sum_{\nu=1}^{j+1} G_\nu^{(j)} \sin \nu\sigma\tau,$$
$$r_j = \sum_{\nu=0}^{j+1} H_\nu^{(j)} \cos \nu\sigma\tau, \qquad (j = 0, \cdots, n-1),$$

where the $F_\nu^{(j)}$, $G_\nu^{(j)}$, $H_\nu^{(j)}$ are functions of m. Further, let us suppose that $\delta_1, \cdots, \delta_{n-1}$ have been uniquely determined. We wish to show from these assumptions, from the differential equations, and from the initial and periodicity conditions that p_n, q_n, r_n have the same form as (33) for $j = n$, and that δ_n is a uniquely determined constant.

The differential equations obtained by equating the coefficients of γ^n in (13′) are

$$[D^2 + 3(1-m^2)]p_n - 3m(1-m^2)^{\frac{1}{2}}r_n = P^{(n)},$$
$$(34) \qquad -3m(1-m^2)^{\frac{1}{2}}p_n + [D^2 + 1 + 3m^2]r_n = R^{(n)} + 2C_1^{(n)},$$
$$\dot{q}_n = -2rn + \int Q^{(n)} d\tau + C_1^{(n)} - \delta_n r_0,$$

where

$$
\begin{aligned}
P^{(n)} &= -3\delta_n(1-m^2)p_0 + 3\delta_n m(1-m^2)^{1/2}r_0 \\
&\qquad + \text{terms in } p_j,\ q_j,\ r_j,\ \delta_j, \\
R^{(n)} &= 3\delta_n m(1-m^2)p_0 - \delta_n(1+3m^2)r_0 \\
&\qquad + \text{terms in } p_j,\ q_j,\ r_j,\ \delta_j, \\
Q^{(n)} &= \text{terms in } p_j,\ q_j,\ r_j,\ \delta_j, \qquad (j=0,\cdots,n-1;\ \delta_0=0).
\end{aligned}
$$

The undetermined constant δ_n enters the right members only where it is expressed and not in the other terms. In $P^{(n)}$ and $R^{(n)}$ the powers of the q's are even while in $Q^{(n)}$ they are odd. Hence $P^{(n)}$ and $R^{(n)}$ are sums of cosines of multiples of $\sigma\tau$ while $Q^{(n)}$ is a sum of sines of multiples of $\sigma\tau$. They have the form

$$
\begin{aligned}
P^{(n)} &= a_0^{(n)} + (d_1^{(n)}\delta_n + a_1^{(n)})\cos\sigma\tau + \cdots + a^{(n)}_{n+1}\cos(n+1)\sigma\tau, \\
R^{(n)} &= c_0^{(n)} + (d_2^{(n)}\delta_n + c_1^{(n)})\cos\sigma\tau + \cdots + c^{(n)}_{n+1}\cos(n+1)\sigma\tau, \\
Q^{(n)} &= b_1^{(n)}\sin\sigma\tau + \cdots + b^{(n)}_{n+1}\sin(n+1)\sigma\tau,
\end{aligned}
$$

The complementary functions of (34, a and b) and the terms arising from $2C_1^{(2)}$ in (34 b) are

$$
\begin{aligned}
p_n &= \omega(B_1^{(n)}e^{i\sigma\tau} + B_2^{(n)}e^{-i\sigma\tau}) + 2m(1-m^2)^{-1/2}C_1^{(n)}, \\
r_n &= B_1^{(n)}e^{i\sigma\tau} + B_2^{(n)}e^{-i\sigma\tau} + 2C_1^{(n)}.
\end{aligned}
$$

The symbolic expressions for the particular integrals are the same as (25) with the appropriate changes in subscripts and superscripts. As at the previous steps the coefficients of $\cos\sigma\tau$ in the numerators of these expressions must vanish in order that p_n and q_n shall be periodic. We thus arrive at the two equations

$$
\begin{aligned}
(1-\sigma^2+3m^2)(d_1^{(n)}\delta_n + a_1^{(n)}) + 3m(1-m^2)^{1/2}(d_2^{(n)}\delta_n + c_1^{(n)}) &= 0, \\
3m(1-m^2)^{1/2}(d_1^{(n)}\delta_n + a_1^{(n)}) + [-\sigma^2+3(1-m^2)](d_2^{(n)}\delta_n + c_1^{(n)}) &= 0.
\end{aligned}
$$

Since the functional determinant in these equations vanishes, the two equations are equivalent and can be satisfied by solving either for δ_n. Thus

$$
\delta_n = \frac{-(-\sigma^2+1+3m^2)a_1^{(n)} - 3m(1-m^2)^{1/2}c_1^{(n)}}{(-\sigma^2+1+3m^2)d_1^{(n)} + 3m(1-m^2)^{1/2}d_2^{(n)}}.
$$

With this choice of δ_n the particular integrals will be periodic and will have the form

$$
\begin{aligned}
p_n &= \alpha_0^{(n)} + \alpha_2^{(n)}\cos 2\sigma\tau + \cdots + \alpha^{(n)}_{n+1}\cos(n+1)\sigma\tau, \\
r_n &= \gamma_0^{(n)} + \gamma_2^{(n)}\cos 2\sigma\tau + \cdots + \gamma^{(n)}_{n+1}\cos(n+1)\sigma\tau,
\end{aligned}
$$

On substituting the complete solution for r_n in (34 c) and integrating we obtain

$$q_n = (2i/\sigma)(B_1^{(n)}e^{i\sigma\tau} - B_2^{(n)}e^{-i\sigma\tau}) - (3C_1^{(n)} + 2\gamma_0^{(n)})\tau$$
$$+ C_2^{(n)} + \sum_{\nu=2}^{n+1} \beta_\nu^{(n)} \sin \nu\sigma\tau,$$

and in order that this solution shall be periodic we must put

$$C_1^{(n)} = -(2/3)\gamma_0^{(n)}.$$

When the initial conditions are applied we obtain

$$C_2^{(n)} = 0, \quad B_1^{(n)} = B_2^{(n)} = \text{a constant}.$$

Hence p_n, q_n and r_n have the same form as (33) when $j = n$. This completes the induction. The construction of the solutions can therefore be carried on to any desired degree of accuracy.

The two sets of solutions can be obtained by restoring the subscripts 1 or 2 to ω and σ.

9. *The Final Form of the Solutions.* On substituting the various values for p_j, q_j, r_j in (19) and the results in (9) we obtain

$$\begin{aligned}
x &= m + \sum_{j=0}^{\infty} \Big(\sum_{\nu=0}^{j+1} F_\nu^{(j)} \cos \nu\sigma\tau\Big)\gamma^{j+1}, \\
\eta &= 0 + \sum_{j=0}^{\infty} \Big(\sum_{\nu=1}^{j+1} G_\nu^{(j)} \sin \nu\sigma\tau\Big)\gamma^{j+1}, \\
\xi &= (1 - m^2)^{1/2} + \sum_{j=0}^{\infty} \Big(\sum_{\nu=0}^{j+1} H_\nu^{(j)} \cos \nu\sigma\tau\Big)\gamma^{j+1}, \\
\tau &= (1 + \sum_{j=1}^{\infty} \delta_j\gamma^j)^{-1/2}(t - t_0).
\end{aligned}$$

In the above equations m, γ and t_0 are the only parameters which remain arbitrary; m denoting the scale factor of the circular orbits, γ that of the periodic oscillations near these orbits, and t_0 the epoch. By substituting for η and ξ in the equations

$$\begin{aligned}
y &= -\eta \cos (t - t_0) + \xi \sin (t - t_0), \\
z &= \eta \sin (t - t_0) + \xi \cos (t - t_0),
\end{aligned}$$

we may obtain the corresponding values of y and z. There are two sets of values, x_1, y_1, z_1; x_2, y_2, z_2, corresponding to the two electrons, but they are not independent inasmuch as the restrictions (2) hold

$$x_1 = -x_2, \quad y_1 = y_2, \quad z_1 = z_2.$$

10. *Numerical Example.* Mr. Smith assigned the values

$$k^2 = .5, \quad m = .5, \quad \gamma = .05, \quad t_0 = 0,$$

and on completing the integrations up to the terms in p_2, q_2 and r_2 he obtained

$$p = -.0025 + .064 \cos \sigma\tau + .025 \cos 2\sigma\tau - .002 \cos 3\sigma\tau,$$
$$q = -.19 \sin \sigma\tau + .03 \sin 2\sigma\tau - .0003 \sin 3\sigma\tau,$$
$$r = .0043 + .077 \cos \sigma\tau - .017 \cos 2\sigma\tau - .00027 \cos 3\sigma\tau.$$

Using the subscript 1 on ω and σ he found

$$\sigma_1 = .825, \quad P_1 = 12\pi/5, \quad \text{nearly.}$$

Values of t were then taken at approximately 30° intervals as t ranges from 0° to 2160°, that is, through the complete period, and the numerical values of x_1, y_1, and z_1 were computed. The values obtained near the beginning and near the end of the period are found in the accompanying Table.

t^0	x_1	y_1	z_1
0	.560	.00	.95
30	.555	.54	.77
60	.536	.87	.33
90	.515	.87	$-.19$
120	.488	.64	$-.59$
150	.464	.26	$-.78$
180	.450	$-.10$	$-.80$
210	.440	$-.43$	$-.66$
240	.475	$-.65$	$-.46$
270	.458	$-.81$	$-.14$
300	.480	$-.81$	.26
330	.506	$-.60$	.65
360	.530	$-.17$	.83
. .	. .	. .	. .
. .	. .	. .	. .
1800	.530	.17	.93
1836	.500	.68	.57
1890	.458	.81	$-.14$
1926	.443	.60	$-.51$
1980	.448	.10	$-.80$
2016	.470	$-.36$	$-.76$
2070	.515	$-.87$	$-.19$
2106	.542	$-.81$	.43
2160	.560	0	.95

A check was made on the work by making use of the *vis viva* integral (4 a). Various sets of computed values for x_1, y_1, z_1 and their derivatives were used and the constant in the *vis viva* integral was found to range from 2.17 to 2.31.

The accompanying diagrams give the projections of the oscillations on the coördinate planes. The circular orbit is not shown in Fig. 2. Its projections in Fig. 1 and Fig. 3 are the y- and z-axes respectively.

11. *Two-Dimensional Orbits.* Two-dimensional periodic oscillations near the circular orbits can be readily found by neglecting the terms in x in the preceding construction. These orbits are coplanar with the circular orbits.

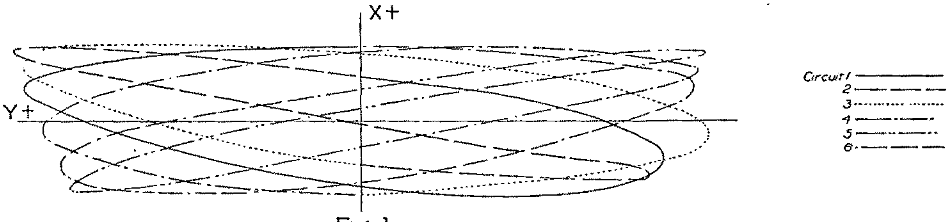


Fig. 1

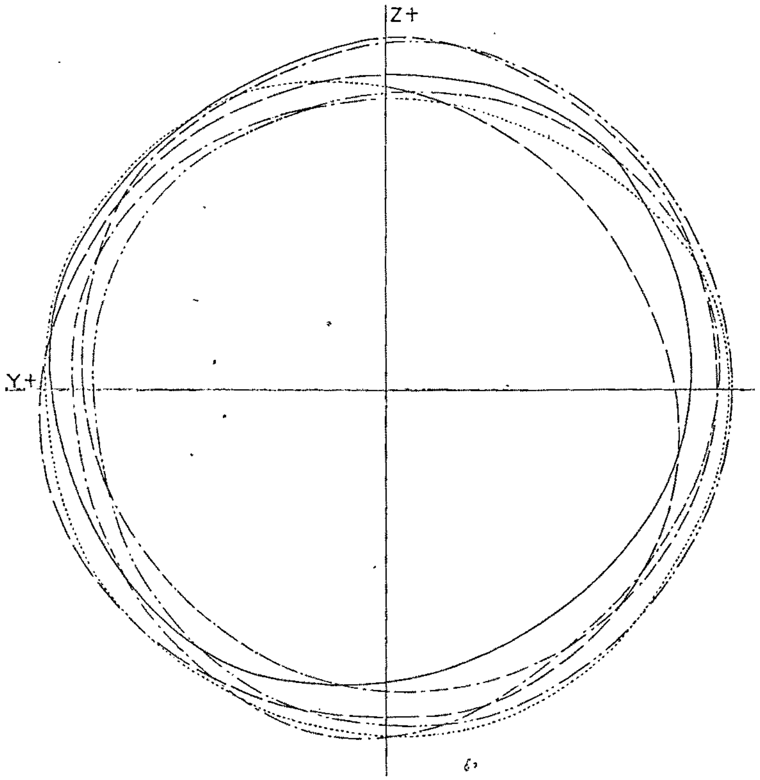


Fig. 2

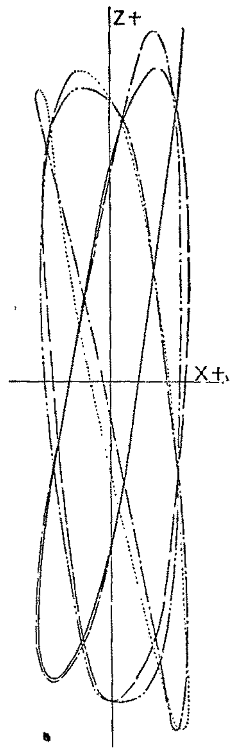


Fig. 3

The actual construction was carried out but as no peculiarities were found it is omitted. Mr. Smith computed an orbit and found curves similar to those in Fig. 2.

The University of British Columbia,
Vancouver, Canada.

On the Groups Which Contain a Given Invariant Subgroup and Transform It According to a Given Operator in Its Group of Isomorphisms.

By H. R. BRAHANA.

A method by which one may construct all the groups which contain a given group H as an invariant subgroup of prime index p was given recently by Professor Miller.* In the papers cited the method was applied and several theorems were introduced which accomplished simplifications of the method in special cases, mostly cases in which H was abelian or the isomorphism performed on H by an operator outside H was of order p. The subject was presented by Professor Miller to a class which the writer attended and after discussion it was decided to investigate the possible wider application of these theorems. The results of this investigation are offered here.

We consider a group H and a group G of order $p \cdot h$ which contains H as an invariant subgroup of prime index p. Let t_1 be an operator outside H. Its p-th power will be in H, and G is generated by t_1 and H. Following the method used by Miller (*loc. cit.*) we may consider G to be written as a regular group in which H is intransitive but is transitive on the h letters of each of p constituents. The operator t_1 permutes these constituents cyclically. Let t be an operator on the $p \cdot h$ letters which permutes the transitive constituents of H in the same way as t_1, but which transforms every operator of H into itself. Then the operator $t_1 t^{-1}$ will transform each of the transitive constituents into itself and will transform the operators of H in the same way as t_1. Let $t_1 t^{-1} = s_1' s_2' \cdots s_p'$, where s_i' is that part of the product $t_1 t^{-1}$ which involves only letters of the i-th constituent H_i. s_1' transforms H_1 in the same way as some operator s_1 in its group of isomorphisms. Let us define $s_2, s_3, \cdots, s_p$ by the relation $t^{-1} s_i t = s_{i+1}$. Then t_1 which is $s_1' s_2' \cdots s_p' t$ performs the same transformation on H as $s_1 s_2 \cdots s_p t$. The operator $Q = t_1^p$ is in H and hence is permutable with t. Q transforms H in the same way as $s_1^p s_2^p \cdots s_p^p$. The operator $Q \cdot s_1^{-p} s_2^{-p} \cdots s_p^{-p}$ which we shall denote by $\bar{s}_0' \bar{s}_0'' \cdots \bar{s}_0^{(p)}$, where $\bar{s}_0^{(i)}$ is that part of the product which involves only

* (1) *Proceedings of the National Academy of Sciences*, Vol. 14 (1928), p. 819. See also (2) *loc. cit.*, p. 918; and (3) *Transactions of the American Mathematical Society*, Vol. 2 (1901), p. 264, and (4) *American Journal of Mathematics*, Vol. 24 (1902), p. 395, in which he described and used the method in the construction of prime power groups.

letters of the i-th transitive constituent, is permutable with every operator of H and also with t. This operator is in the conjoint of H and moreover it is transformed into itself by s_i since this is true of both Q and s_j. Now let us consider the operator $U = \bar{s}_0' s_1 s_2 \cdots s_p t$. U transforms H in the same manner as t_1 and its p-th power is $\bar{s}_0' \bar{s}_0'' \cdots \bar{s}_0^{(p)} s_1^p s_2^p \cdots s_p^p$ which is Q. Therefore, $\{H, U\}$ is simply isomorphic with $\{H, t_1\}$.

Conversely, if there exists an operator s in the group of isomorphisms of H whose p-th power is an inner isomorphism and an operator Q in H which transforms the operators of H in the same way as s^p and is invariant under s, then the operator $\bar{s}_0'$ and consequently the operator U and the group G exist. Therefore,

A necessary and sufficient condition that there exists a group G of order $p \cdot h$ in which the operators of a given invariant subgroup H are transformed according to an operator s in its group of isomorphisms whose p-th power is an inner isomorphism is that there exists an operator Q of H which transforms the operators of H according to s^p and is invariant under s.

The operator $\bar{s}_0'$, and consequently U also, is completely determined by Q and s. s does not determine Q completely but determines it as one of a set of operators of H each of which transforms H in the same manner as s^p and each of which is permutable with s. The operators of H which transform H in the same manner as s^p may all be obtained from one of them by multiplying it in turn by operators from the central of H. The operators of the central of H which are permutable with s form a subgroup C which when s is not identity is the central of G and does not depend on Q. Therefore,

Every group G which contains a given group H invariantly as a subgroup of prime index p and transforms it according to a given operator s, not identity, in its group of isomorphisms contains a central C which depends only on H and s.

The order of s is of necessity a multiple of p, but in any group G the operator U may be so chosen that it transforms H according to an operator s whose order is a power of p, for if the order of the transformation performed by U is $m \cdot p^a$ where m is prime to p then U^m will transform H according to an operator s whose order is a power of p. The groups $\{H, U\}$ and $\{H, U^m\}$ are evidently the same. We shall therefore assume in what follows that the order of s is a power of p.

A necessary and sufficient condition that for a given H and s there exist a group G is given in the first theorem. That such a group need not always exist was shown by Professor Miller.* We shall accordingly in what follows

* *loc. cit.*, (1) p. 821.

assume that one such group exists for the H and s under consideration and investigate the question of the existence of other groups.

If the given group is $\{H, U\}$ where $U^p = Q$, every possible group determined by H and s is generated by H and an operator which transforms H according to s and which has $C_i \cdot Q$ for a p-th power, where C_i is some operator of C. Since s is of order p^a, $U^{p^a} = R \cdot Q'$, where both R and Q' are in C, the order of R is prime to p, and the order of Q' is a power of p. Therefore, the group $\{H, U\}$ will contain an operator $\bar{U} = R^k \cdot U$ which transforms the operators of H according to s and whose p^a-th power is Q' of order a power of p. Since the groups $\{H, U\}$, and $\{H, \bar{U}\}$ are the same, we may assume that the order of Q is a power of p.

Now any other group that corresponds to H and s may be obtained by taking H and $s_0 U$ where s_0 is chosen so that $s_0 s_0' \cdot \cdot \cdot s_0^{(p)}$ is an operator in C; and though every operator of C will give an s_0 and every s_0 determines a group, it follows from the preceding paragraph that the number of distinct groups cannot exceed the order of the Sylow subgroup of order p^γ in C.

If R is any operator of C then $(RU)^p = R^p Q$. Therefore, the group obtained by taking s_0 to correspond to R^p is the same as that obtained by taking s_0 to be identity. We have the theorem:

The number of groups which contain a given group H invariantly as a subgroup of index p and transform its operators according to a given operator s in its group of isomorphisms is not more than one greater than the number of operators which are not p-th powers in the Sylow subgroup of order p^γ in the central C.

An operator of G which transforms H in the same manner as U must be the product of U and an operator from the central of H, and if it has for a p-th power the product of Q and an operator from the Sylow subgroup C_{p^γ} of order p^γ of C the operator from the central of H must be from its Sylow subgroup H_{p^β} of order p^β. Let R be such an operator, let $U^{-1}RU = R_1 R$, and let $U^{-1}R_i U = R_{i+1} R_i$. Then since $U^p = Q$, we have $U^{-p}RU^p = R_p R_{p-1}^{\binom{p}{1}} R_{p-2}^{\binom{p}{2}} \cdot \cdot \cdot R_1 \quad R = R$, where the exponents are the binomial coefficients. From this we get

$$R_p R_{p-1}^{\binom{p}{1}} R_{p-2}^{\binom{p}{2}} \cdot \cdot \cdot R_1^{\binom{p}{1}} = 1. \tag{1}$$

Then $(RU)^p = U^p \cdot R_p R_{p-1}^{\binom{p+1}{1}} \cdot \cdot \cdot R_1^{\binom{p+1}{2}} \cdot R^p$, which in view of (1) becomes

$$(RU)^p = Q \cdot R_{p-1} R_{p-2}^{\binom{p}{1}} R_{p-3}^{\binom{p}{2}} \cdot \cdot \cdot R_1^{\binom{p}{2}} \cdot R^p. \tag{2}$$

If R' is another operator in the central of H the operator $(R'U)^p$ will be the same as the right member of (2) where the R_i is replaced by R_i'. Then $(R'RU)^p = Q \cdot (R'_{p-1}R_{p-1}) \; (R'_{p-2}R_{p-2})^{\binom{p}{1}} \cdots (R_1'R_1)^{\binom{p}{2}} (R'R)^p$ $= Q \cdot R'_{p-1}R'^{\binom{p}{1}}_{p-2} \cdots R'^{\binom{p}{2}}_1 R'^p \cdot R_{p-1}R^{\binom{p}{1}}_{p-2} \cdots R^{\binom{p}{2}}_1 R^p$. Hence the operators of $C_p{}^\gamma$ which with Q determine p-th powers of operators of G which transform H in the same manner as U form a group; we shall denote this group by C_M. Moreover, the set of operators $R_{p-1}R^{\binom{p}{1}}_{p-2} R^{\binom{p}{2}}_{p-3} \cdots R^{\binom{p}{2}}_1 R^p$, where R is allowed to go through a set of independent generators of the Sylow subgroup of order p^β of the central of H generate a group which contains every operator in the central of H which can be written in that form. The cross-cut of this group and C is C_M.

If $C_p{}^\gamma$ is arranged in co-sets with respect to C_M, a choice of s_0 which makes the product of Q and one operator of a particular co-set the p-th power of an operator which transforms H in the same manner as U, makes the product of Q and every operator of that co-set such a p-th power. Therefore,

The number of groups determined by a given H and s does not exceed the order of the quotient group of $C_p{}^\gamma$ with respect to C_M.

It is true that we may determine an s_0 for each operator of the quotient group of $C_p{}^\gamma$ with respect to C_M and that each such s_0 determines a group G which has a new set of operators for p-th powers of operators which transform H in the same way as U, but we may not conclude therefrom that there are that many distinct groups G, due to the possibility of isomorphisms of H which are permutable with s. This will become more apparent when we consider certain restrictions on H and s.

The method of procedure indicated in the proof of the preceding theorem is quite readily carried out when both p and the number of invariants of $C_p{}^\gamma$ are small. Often, however, the result may be arrived at indirectly in a simpler manner. From the form of the right member of (2) we notice that every operator of C_M is in the group H_M generated by the p-th powers of operators in the Sylow subgroup of order p^β in the central of H and the $(p-1)$-th derived group of this Sylow subgroup with respect to U. Since C_M is in C, C_M will be in the cross-cut of H_M and C; we shall denote this cross-cut by C_L.

We shall now consider some of the subgroups of C_M. In any case where we can show that such a subgroup coincides with C_L, we may conclude that C_M coincides with this subgroup.

Let us consider an operator R in the central of H whose p-th power is in C_p^{γ}. Then $U^{-1}R^pU = R_1^pR^p$ which must be R^p. Therefore, R_1 and each of the succeeding R_i's must be of order p or 1. Then from (1) we see that R_p must be identity, which requires R_{p-1} to be in C_p^{γ}. Moreover, (2) reduces to $(RU)^p = Q \cdot R_{p-1}R^p$. If R_{p-1} is identity then the operator R^p is in C_M.* The R's for which the corresponding R_{p-1}'s are identity form a group and their p-th powers form a group which is in C_M and which we shall denote by C_I.

If one of the operators R_{p-1} above is the p-th power of an operator S in C_p^{γ}, then $(S^{-1}RU)^p = Q \cdot S^{-p}R_{p-1}R^p = Q \cdot R^p$. Then R^p is in C_M. The product of two such R's fulfills the same conditions, as do the operators R which determine C_I. Thus we have determined a group C_J which is contained in C_M and contains C_I.

The $(p-1)$-th derived group with respect to U of the set of operators of the central of H whose p-th powers are in C_p^{γ} is, as we have seen, contained in C_p^{γ} and is of type $1, 1, \cdots$. The group C_I contains all of those R_{p-1}'s which are p-th powers in C_p^{γ}. For each of the independent generators of the group of R_{p-1}'s which are not in C_I we may determine an operator $R_{p-1}R^p$, any one of which is obtained from a given one by multiplying the latter by some operator from C_I. The group C_K determined by these operators and C_J is contained in C_M and contains C_I.

These three groups may be described as follows: C_I is composed of the set of p-th powers of the set of operators in the central of H whose p-th powers are in C_p^{γ} and whose $(p-1)$-th commutators are identity; C_J is obtained by removing the restriction that the $(p-1)$-th commutators be identity and requiring that they be p-th powers in C_p^{γ}; and C_K is obtained by extending C_J by means of a definite operator for each of the remaining generators of the $(p-1)$-th derived group of the set of operators in the central of H whose p-th powers are in C_p^{γ}.

If we suppose that R is an operator in the group H_p^{β} for which R_{p-1} and $R_{p-1}R_{p-2}^{\binom{p}{1}} \cdots R_1^{\binom{p}{2}} R^p$ are in C_p^{γ}, we note first that R_{p-1} is of order p, since $U^{-p}R_{p-2}U^p = R_{p-1}^{\binom{p}{1}} R^1_{p-2} = R_{p-2}$. Then from $U^{-p}R_{p-3}U^p = R_{p-1}^{\binom{p}{2}} R_{p-2}^{\binom{p}{1}} R_{p-3} = R_{p-3}$ it follows that R_{p-2} is also of order p. By repetition of this process we may show that every R_i is of order p, and that therefore $R_{p-1}R_{p-2}^{\binom{p}{1}} \cdots R_1^{\binom{p}{2}} R^p$ becomes $R_{p-1}R^p$. Hence under the conditions on R its p-th power must be in

* This includes two of Miller's theorems: (1) R is in C, *loc. cit.* (1), p. 820; and (2) R_1 is in C, is of order p, *loc. cit.* (3), p. 265.

C_p^γ and the subgroup C_K of C_M cannot be extended by an operator corr sponding to an R_{p-1} which is invariant.

To continue to a consideration of the R_{p-1}'s which are non-invariant woul be to give a complete determination of C_M for which a method has alreac been pointed out. From the foregoing a number of conclusions concernii special cases may be drawn; we shall give three.

(a) *If the $(p-1)$-th derived group of H_p^β is identity, then C_M coi cides with C_I; if it is composed of p-th powers in C_p^γ, then C_M coi cides with C_J; if it is composed of invariant operators, then C coincides with C_K.*

(b) *If the group of p-th powers of operators of H_p^β is contained C_p^γ, then the $(p-1)$-th derived group of H_p^β is in C_p^γ and C coincides with C_K.*

(c) *If C_p^γ coincides with the group of p-th powers of operators of H_t and if the $(p-1)$-th derived group of H_p^β is contained in the grou of p-th powers of operators of C_p^γ, then C_M coincides with C_J ar with C_p^γ, and therefore there is but one group corresponding H and s.*

If R is an operator of H_p^β which is transformed into its k-th pow by U, then $(RU)^p = Q \cdot R^{1+k+k^2+\cdots+k^{p-1}}$. If this operator $R^{1+k+\cdots+k^{p-1}}$ is C_p^γ it is in C_M.* If we have determined C_M this gives us no new informatio but if we are determining C_I, C_J, or C_K it gives additional information co cerning C_M whenever $R^{1+k+\cdots+k^{p-1}}$ is not in C_K.

Thus far we have placed no restrictions on H or s. Let us now suppo that s is of order p. The operator s^p is permutable with every operator of I Since H always contains at least one operator, namely identity, which is i variant under H and s, Q always exists. Therefore,

For a given H and an s of order p in its group of isomorphisms the exists at least one group G which contains H invariantly and transforms according to s.

If H is abelian s must be of order p. This makes no change in the pr cedure in the determination of C_M since that depended only on the operato of H_p^β, which were permutable with each other and with s^p. However, wh H is abelian it is the direct product of its Sylow subgroups and its group isomorphisms is the direct product of the groups of isomorphisms of its Sylc

* This theorem is given by Miller for H abelian, *loc. cit.* (2), p. 918.

subgroups. These Sylow subgroups are abelian and therefore the group of isomorphisms of H ... invariant operators which permute [illegible] phisms on the Sylow subgroups. Hence when H is abelian the group obtained by extending H by means of s_0 is simply isomorphic with that obtained by extending H by means of $s_0^k H$ where k is prime to p, and therefore in the determination of s_0 it is necessary to consider but one operator, and that of highest order, from any cyclic subgroup. Hence,

If H is abelian the number of groups G which contain H invariantly and transform it according to a given operator of order p in its group of isomorphisms is not more than one greater than the number of cyclic groups which are not contained in cyclic groups of higher order of the quotient group of C_p^γ with respect to C_H.

If H is cyclic all the above groups are cyclic and therefore there cannot be more than two groups for a given s.[*] If H is cyclic and s does not leave invariant the operators of highest order of H_p^γ, then C_H coincides with C_p^γ except that when $p = 2$ and C_2^γ is of order 2 then we have $R \cdot R^2 = 1$ where R is the operator of order 4. Thus when H is cyclic there are two groups only if s leaves the operators of highest order of H_p^β invariant, or, when $p = 2$, transforms them into their inverses.

The theorem just stated for H abelian is not true when H is non-abelian and we shall conclude by giving an example to prove it.

Let H be $\{s_1, s_2, s_3\}$ where s_1, s_2, and s_3 satisfy the conditions

$$(1) \qquad \begin{cases} s_1^p = 1, \quad s_2^p = 1, \quad s_1 s_2 = [illegible] \\ s_3^p = s_1, \quad s_3^{-1} s_2 s_3 = s_1 s_2. \end{cases}$$

It is obvious that H is a non-abelian group of order p^3 and contains an abelian subgroup of order p^2 and type 1, 1. Now let us consider the groups G and $\bar{G}$ obtained by adjoining operators s_4 and s_4 which satisfy respectively the relations

$$(2) \qquad s_4^{-1} s_2 s_4 = s_1 s_2, \quad s_4^{-1} s_3 s_4 = s_2 s_3, \quad s_4^p = s_1,$$

and

$$(\bar{2}) \qquad s_4^{-1} s_2 s_4 = s_1 s_2, \quad s_4^{-1} s_3 s_4 = s_2 s_3, \quad s_4^p = s_1^2.$$

The groups G and $\bar{G}$ are both of order p^4, the operators s_4 ... perform the same transformation on operators of H and are of the same order. In

---

[*] Burnside, *Theory of Groups*, 1911 ... [illegible] ... 1.2 ... p. 20 ...

[†] G. A. Miller, *loc. cit.* [illegible]